UNIVERSITY OF CANTERBURY

MASTER THESIS

Augmented Virtuality Enhanced Visualization in an Immersive Cinematic Environment

Author:
Wenjing TANG

Supervisor:
Prof. Rob LINDEMAN

Co-Supervisor:
Dr. Gun LEE

Associate Supervisor:
Prof. Mark BILLINGHURST

A thesis submitted in fulfillment of the requirements for the degree of Master of Human Interface Technology

in the

HIT Lab NZ
College of Engineering

April 2018

Declaration of Authorship

I, WENJING TANG, declare that this thesis titled, 'Augmented Virtuality Enhanced Visualization in an Immersive Cinematic Environment' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: Wenjing Tang

Date: 20/04/2018

UNIVERSITY OF CANTERBURY

Abstract

Human Interface Technology Lab New Zealand
Department of Engineering

Master of Human Interface Technology

by Wenjing TANG

The rapid development of affordable head-mounted displays (HMDs) has led to inclusion of Virtual Reality (VR) in a home entertainment system, which in turn has created a niche for 360 degree panoramic movies. Previous research blended a user and real-world objects into a virtual movie scene seamlessly, making the user feel being part of the virtual environment as if they were in the movie space. This thesis further developed the concept by overlaying context-aware virtual costumes on the user's real body. A prototype was developed by combining Microsoft Kinect, a SoftKinetic depth camera, and a HMD. The prototype captured user's real body and embedded it in a virtual 360 movie scene; augmenting the virtual scene with reality resulting in augmented virtuality (AV). Furthermore, virtual costumes related to the movie scene were overlaid on user's real body to enhance user experience. The virtual content, captured real body, and 360 degree movie were combined in Unity and visualized in the Oculus DK2 HMD. With the created prototype, a user experiment was conducted to investigate how context-aware virtual costumes on user's real body affected the user's sense of presence and preference in a 360° movie scene. Results showed augmenting user's real body with context-aware virtual costumes was most preferred by users, compared to only watching a movie and just augmenting user's real body. The results offer a future direction to generate greater enhanced 360 movie watching experiences in a HMD.

Acknowledgements

I would like to express sincere gratitude to my senior supervisor Prof. Lindeman and co-supervisor Dr. Lee for their guidance and supervision throughout this work. This thesis would not have been possible without their support, both valuable bits of advice for the project and insightful feedback for the thesis. Many thanks to all HITLabNZ students and staff especially Humayun Khan, Qi Min Ser, Omprakash Rudhru, Jackie Soady and Ken Beckman for their support and encouragement through the tough times. A special thank to Wenyi for his help, assistance and support in the aspects of technology and life. Finally, I would like to thank my family for their care, encouragement and support, as well as Faye and Rayne for their love and sacrifice.

Contents

List of Figures

List of Tables

Abbreviations

2D	**T**wo-**D**imensional
3D	**T**hree-**D**imensional
4D	**F**our-**D**imensional
360°	**3**60-**D**egree **P**anoramic
VR	**V**irtual **R**eality
HMD	**H**ead-**M**ounted **D**isplay
MR	**M**ixed **R**eality
AV	**A**ugmented **V**irtuality
AR	**A**ugmented **R**eality
SDK	**S**oftware **D**evelopment **K**it
DK	**D**evelopment **K**it
FPS	**F**rame **P**er **S**econd
CGI	**C**omputer-Generated **I**magery
RGB	**R**ed **G**reen **B**lue
GPU	**G**raphic **P**rocessing **U**nit
HCI	**H**uman **C**omputer **I**nteraction
AI	**A**rtificial **I**ntelligence
API	**A**pplication **P**rogrammming **I**nterface
DLL	**D**ynamic **L**inked **L**ibrary
NUI	**N**ational-**U**ser **I**nterface
IPQ	**I**group **P**resence **Q**uestionnaire
PRES	**G**eneral **P**resence
SP	**S**patial **P**resence
INV	**I**nvolvement
REAL	**E**xperienced **R**ealism

Dedicated to Faye and Rayne . . .

Chapter 1

Introduction

1.1 Background and Motivation

Since the birth of films, for more than 100 years filmmakers have been working from a technical perspective to enhance the audience's viewing experience, from silence to sound, from black and white to colour, and from two-dimensional (2D) to three-dimensions (3D). The recent development of computer technologies has allowed the way of displaying a film to evolve from 3D further to four-dimensional (4D). A 4D film entails a computer system and sensor technologies that are added to the 3D vision system, and it is a new human-computer interaction experience that is mixed with various simulated special effects, including visual, auditory, tactile, and olfactory [1]. The high-tech Virtual Reality (VR) technology is bound to have an impact on the traditional film industry. VR technology has revolutionized the way people watch movies. Recent rapid development of the latest high-quality head-mounted displays (HMDs) [1] [2] [3] make immersive VR experiences easily affordable by the general public. It brings immersive cinematic experience from public 4D theatres into a home entertainment system. The development of HMDs has led to the rising popularity of 4D movies in a home entertainment environment. A richer VR experience has been developed in a new market of 360° movies.[4] [5] The 360° movie is a spherical video that contains a panoramic view of the scene. It transforms the ways users communicate, create, collaborate, and explore, which allows users to become immersed in more than just a single view. Currently, a large population looks forward to experiencing VR 360° movies. The HMDs allow viewers to watch movies from any angle without any restrictions. Viewers are free to choose their perspectives, and the

[1]https://www.oculus.com/
[2]https://www.vive.com/nz/
[3]http://www.samsung.com/global/galaxy/gear-vr/
[4]https://www.youtube.com/channel/UCnO5ygba3vsJG9IeZLndaTQ/about
[5]https://facebook360.fb.com/

ability to see the whole scene brings a strong sense of immersiveness and presence. Users will appear to be in the scene or standing next to the character in the film.

Recently, Mixed Reality (MR)[2] has shown great potential and has attracted significant commercial and research interests. MR involves the merging of virtual and real worlds, covering physical reality, augmented reality(AR), augmented virtuality(AV), and virtual reality(VR) [3], as shown in Figure 1.1.

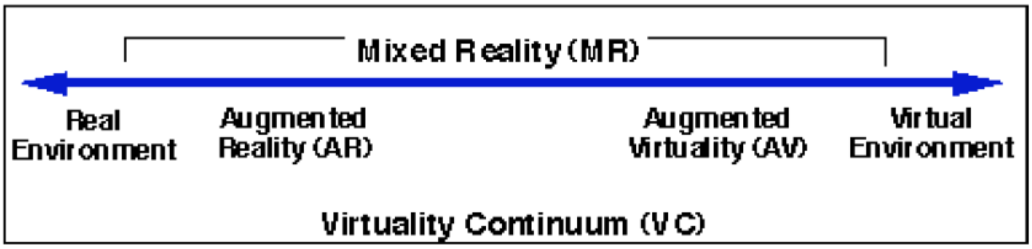


FIGURE 1.1: Simplified representation of a "virtuality continuum" [3]

Stapleton et al. [4] commented that " 'Mixed-reality' technologies combine virtual objects with the real world to suspend disbelief and engage audiences in a rich fantasy experience." VR immerses the user in a simulated world, while AV refers to the merging of real-world objects into virtual worlds. MR enables virtual objects to look and sound like they are part of the physical world by understanding the user's surroundings. A new real-time visual experience is created by the mixture of the physical entities and the virtual objects and presented to users' eyes in real time.

Rather than watching in a public movie theatre, users are more relaxed and can quickly become immersed in the scene of the movie in a home-based environment. The audience can get an immersive and interactive view of the movie tour with various enhanced interactions. Previous research [5] has successfully achieved the blending of real objects around the viewer to create 360° video cinematic scenes. To bring the immersive experience to the next level is to allow users to have a more natural way of interacting with the virtual content on the screen. Users can visualize both digital content and real world objects around them. This thesis explores a prototype consisting of the enhanced AV visualization, and describes a user experiment to analyze whether this enhanced AV visualization would improve the user's sense of presence.

Microsoft Kinect (Kinect 1) was first released in November 2010 along with Microsoft Xbox. In 2011, Microsoft released its Kinect Software Development Kit (SDK), which allowed Kinect to be used in other fields beside games. In 2014, Microsoft released second Kinect for Windows (Kinect 2). There are several technical improvements to Kinect 2 comparing to Kinect 1, such as higher resolution, higher precision in motion tracking, and face recognition. Initially, Kinect 2 was considered to be used in this research,

Figure 1.2: User interaction using a HMD in a 360° video

however, with the restriction of the available hardware device, the researcher decided to use Kinect 1 as the primary device to track real body motion. The world's leading supplier in 3D vision and gesture recognition solutions, SoftKinetic DepthSense 325 has released a depth camera which allows fully engaged users to experience AR with more immersive interactions, less motion sickness, and most realistic hand representation [6]. A combination of the Kinect and Softkinectic offers the possibility to blend real body with virtual objects in the virtual world in a 360° video cinematic scene.

To specify the environment, a system prototype has been set up in a family-friendly home-like space. Users could choose either to sit down or stand while they are watching the movie. Figure 1.2 shows the user's interaction with the 360° video.

1.2 Research Questions

The research tries to answer these research questions using the designed system.

- Does the system improve user's sense of presence by enhanced visualization (visualizing the real body with or without real-time augmented context-aware virtual objects) in comparison to visualizing a movie only in a cinematic scene?

- Does the user have a preference in visualizing real-time blending (nothing, only physical object or real-time rendering with context-aware virtual objects) in a cinematic scene?

1.3 Contribution

The main contributions of this thesis are:

- A novel method to provide hand interaction in 360° movies by combining real body captured by the Softkinetic camera, user's movement tracked by the Kinect and context-aware virtual objects augmented on movie scene in Unity 3D. The combination used Kinect's body tracking and SoftKinetic's hand visualization technologies.

- A user study analyzing the effect of enhanced visualization on sense of presence and user preference between visualizing movie only, visualizing user's real body and visualizing user's real body with context-aware virtual objects when experiencing the 360° video using a HMD.

1.4 Thesis Structure

The structure of the thesis is as follows.

Chapter 2: Discusses the related work that has been done in the previous study.

Chapter 3: Describes the design process of the prototype.

Chapter 4: Explains the implementation of the prototype.

Chapter 5: Describes the user experiment, procedure, and the evaluation process.

Chapter 6: Analyzes the results acquired from the user study,

Chapter 7: Discusses the results and limitations of the study.

Chapter 8: Concludes the thesis and explores future improvements.

Chapter 2

Related Work

This chapter aims to discuss the related work which has been done before, including five different sections in total. The first section is the related work in cinematic experience in MR. The second section talks about the AV experience. The third section is about depth camera Microsoft Kinect. The fourth section explains the 3D modelling. The last section is about game engine.

2.1 MR in the Cinematic Experience

The mixed reality is actively investigated both in an academic field and commercial industry. However, it is expected to be used in the broader area of daily life. The previous research has discussed several different sites integrated with MR technology.

There are a few types of research using the MR technology for the theatre experiences. In 2001, the VR theatre concept was mentioned in a new research field [7]. It combines VR experience with the IMAX theatre to provide audiences more interactive activities. They designed a multi-user interaction VR theatre system which was suitable for crowd audiences. An implementation example of the Kyongju VR theatre is displayed in the paper. Figure 2.1 shows the system configuration.

The authors proposed Audience Interaction via the multi-user interaction in a VR theatre. They also pointed out three issues for this Audience Interaction, interaction devices, mapping challenge and interaction function design [7].

Another work on a VR theatre, called Gyeongju VR theatre, made the system extensible, reconfigurable and scalable based on a framework named NAVER [8]. Figure 2.2 shows the system configuration of Gyeongju VR theatre. It is a distributed system which

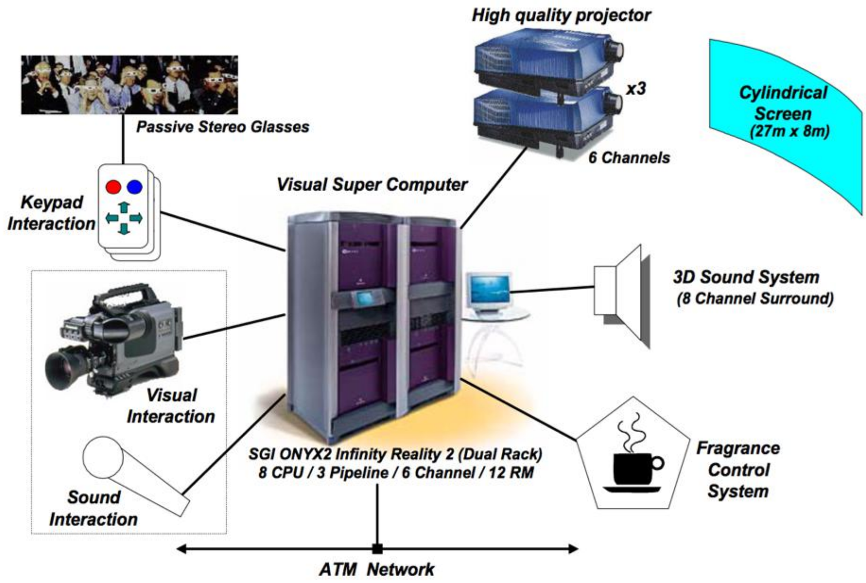


FIGURE 2.1: Kyongju VR theatre system configuration [7]

contained multiple integrated 3D virtual space hosts on the network with a wide range of applications and interfaces.

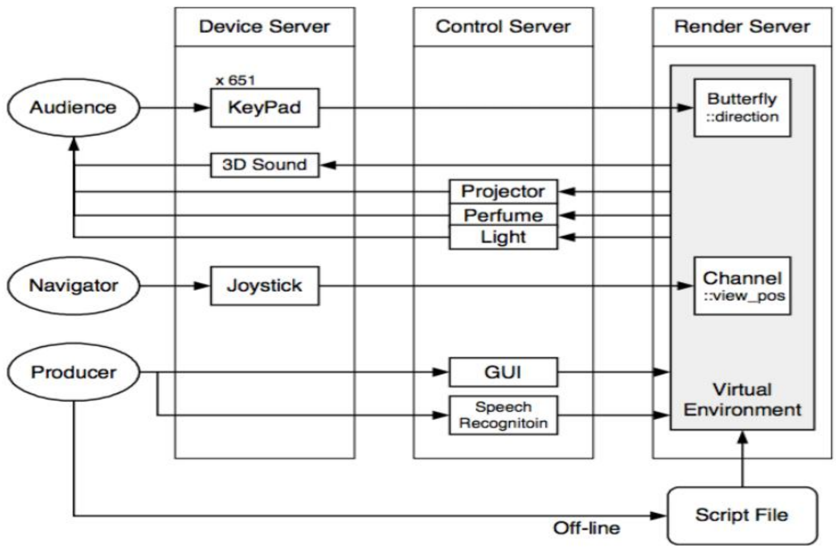


FIGURE 2.2: Gyeongju VR theatre system integration [8]

Another exploration of theatre experiences [9] uses new style MR and wearable computing technology. It allows users to have tangible interaction experiences. Figure 2.3 shows the conceptual diagram of this interactive theatre.

There are three critical stages. Firstly, outdoor theatre land exploration mode which required users to walk around to collect the outdoor environment information with a

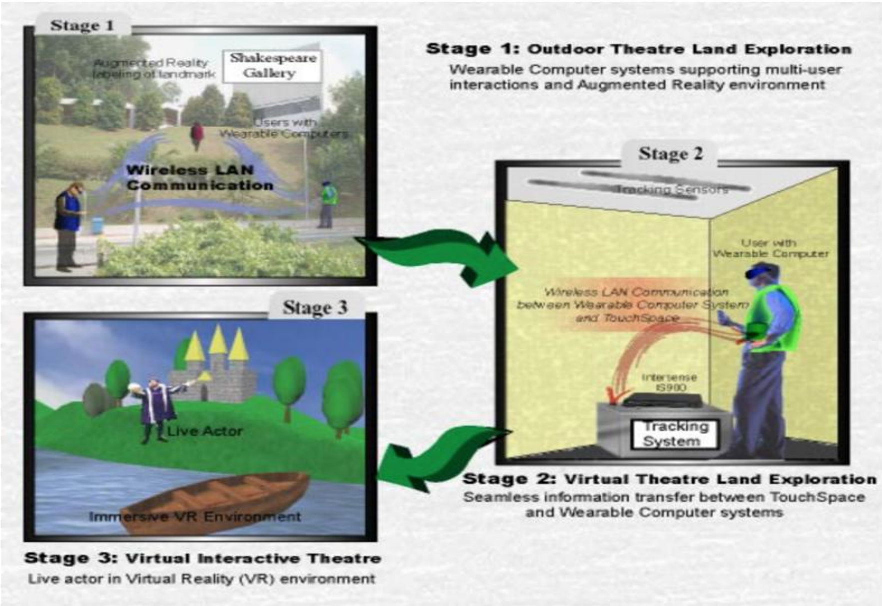


FIGURE 2.3: MR Interactive Theatre idea [9]

wearable computing device. The real environment was overlaid with the virtual environment as AR visualization. Secondly, AR theatre land exploration mode embedded a virtual theatre to the physical environment in a natural way, merging between AR and VR world seamlessly, also allowing interactions with virtual objects and figures seamlessly. Thirdly, the virtual interactive theatre was a fully immersive VR navigator experience mode. It presented a new theatre system which enhanced seamless social interactions in user-to-user and user-to-physical level between real and virtual world.

However, those MR theatre experiences have high requirements for space and hardware which are the obstacles for the users who prefers to enjoy a movie at home with a low cost and time. In comparison, the affordable HMDs offers an easy and cheap way to set up the hardware and space.

With the development of the MR and HMDs devices, some projects of immersive movie experiences have launched. Some VR experience movie projects use the 360° video as well. Framestore studio created interstellar Virtual Reality Experience project [1]. An immersive VR experience was simulated to bring audiences closer to the story.

The Unreal 4 game engine is used. The Interstellar VR Experience allows viewers immerse themselves in an undoubtedly memorable journey through the solar system. This project has attracted crowds to try the VR cinematic experience [10].

[1]https://www.framestore.com/work/interstellar-cinematic-vr-experience

There are some featured films using Gear VR to give viewers a new way to experience movie contents. For example, "The Martian", "The Conjuring 2", "The Jungle Book", and "Star Wars". Table 2.1 lists several films which provide cinematic VR experiences.

Table 2.1: VR cinematic experience movies [11]

Movie	**Device**
Interstellar	Oculus Rift
The Martian	Gear VR
The Conjuring 2	Gear VR
The Jungle Book	Gear VR
Insurgent	Gear VR
Dirrogate	Gear VR
The world inside room	Gear VR
Star Wars	Google cardboard
classic Batman	Samsung's VR content platform

The previous research by Chen [12] has blended user's body into the scene. In this thesis project, the user's body is brought to the scene with context-aware virtual objects.

2.2 Augmented Virtuality

Augmented Virtuality (AV) has been defined in reference to VR environment, in which some "reality" has been added to the immersive VR environment [2]. Researchers aims to enhance user's physical surroundings by integrating virtual objects.

Billinghurst et al. [13] first introduced the concept of transitioning between the virtuality and reality seamlessly. The researchers investigated to how smoothly transport users between reality and virtuality, and demonstrated a prototype system as a proof of concept-magic book interface [14]. It contained three different levels: one was the normal book page which meant the physical object; another was the augmented reality object which was virtual objects that users could see via the display device; and the last one was an immersive virtual space in which users could view virtual avatars in the virtual space. It allowed users to switch between AR viewpoint and VR viewpoint easily. The concept has been explored and extended [15] (shown in Figure 2.4), discussing a transitional interface between AR viewpoint and the VR viewpoint.

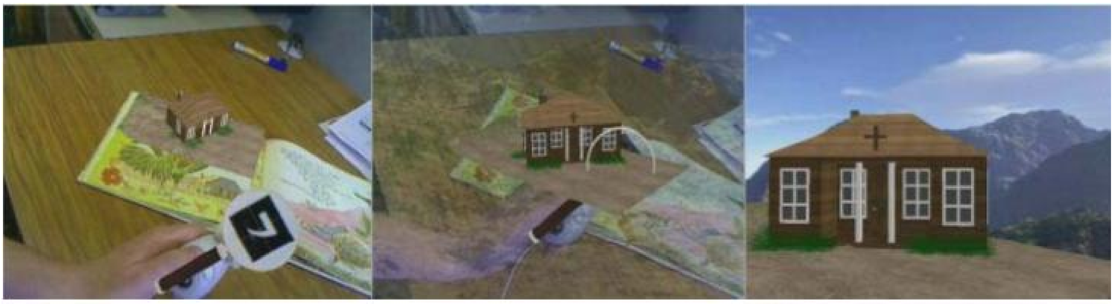


FIGURE 2.4: The example of transitional interface [15]

In 2001, the Mixed Reality Systems Laboratory in Japan created an MR project [11]. They achieved several goals in their research. They created four types of MR visual simulation, as shown in Table 2.2.

TABLE 2.2: Four types of MR visual simulation [11]

Name	Type
Cybercity Walker 2001	a typical AV system
Wisteria World 2001	a telepresence system interacted with MR functions
Seeing Through, Inside Out	a mobile system, and a byproduct of the fourth system
Towards Outdoor Wearable Navigator with Enhanced and Augmented Reality	a pure mobile system

Cybercity Walker 2001 was a typical AV system, which was reconstructed from the real world via walking through and looking around the realistic environment. The authors wanted this to be used in city guide and planning for the cultural value.

Previous research work has created a scene with user body visualization in a cinematic experience [12]. They blended views of the user's body through visualization into an immersive 360 movie.

The system was developed in the Unity game engine. An AV Pro Windows Media plugin was used to transform the 360° spherical panoramic movie into Unity game engine. The Oculus SDK Unity plugin with the head sensor was applied to implement the VR visualization. A Unity plug was developed, using the DepthSense SDK. The Unity game engine called the plugin script to import the depth and colour video stream captured by the SoftKinetic camera. According to the depth perception capacity of SoftKinetic camera, the researcher could supply the Unity game engine with point cloud data, which was obtained by the coordinates x, y, z of points and surfaces directly. Particle system based point cloud rendering was used to generate the complete visualization of physical objects. Figure 2.5 shows the screenshot of their system.

FIGURE 2.5: The screenshot of the system [12]

However, the SoftKinetic camera sensing range is less than 1.5 metres. The images that were more than 1.5 metres away had been shown after the adjustment. The images beyond 1.5 metres were supported by the calibration between the depth and colour sensors to extend the proper mapping. In this thesis project, the SoftKinetic camera sensing range was set to be less than 1.5 metres, because the larger sensing range might affect the sense of immersion in the movie scene with a lot of unnecessary objects.

Both Chen [12] and Khan [16] had worked on hand gesture interaction in a cinematic environment. In Chen's prototype, some built-in gesture feature had been implemented, for example, "Thumbs Up" increased the depth distance while "Thumbs Down" decreased the depth distance.

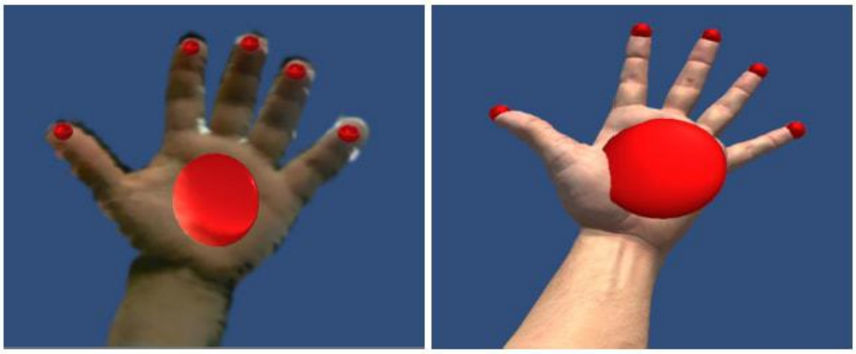


FIGURE 2.6: The corresponding points and hand tracking data of SoftKinetic camera (left) and Leap Motion (Right) [16]

Khan's prototype had an advanced interaction with the user in the cinematic environment. An Oculus DK2 HMD was mounted with SoftKinetic camera and Leap Motion.

DepthSense SDK and Leap Motion Unity 3D asset had been used to calibrate the captured image data in Unity. There was a transformation matrix to align the two datasets from the SoftKinetic camera and Leap Motion together in one space. The real hand texture captured from SoftKinetic camera were overlaid over vitual hand models using the transformation matrix; the combination was rendered in Unity. The correspondence between real and virtual hands is shown in Figure 2.6.

2.3 Microsoft Kinect

Depth sensing camera can provide traditional images and depth images in real-time. Some depth cameras have the Red Green Blue (RGB) camera, while some do not have. This kind of camera has been widely used in AV research.

Kinect is a 3D depth camera which has a wide range of use, and it completely changes the way users experience games and entertainment [17]. Users interact with the systems with their body movements in a natural way. The Kinect sensor consists of the infrared projector, the RGB camera and the infrared camera (Figure 2.7).

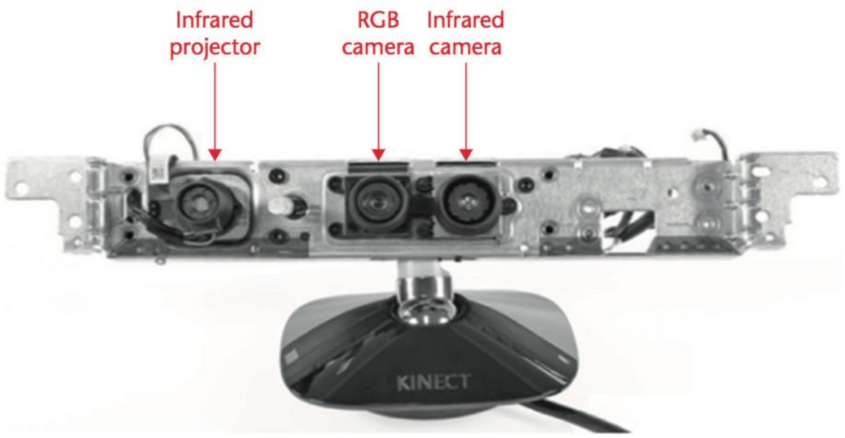


Figure 2.7: Components of the Kinect [17]

The low cost, faster speed and better reliability of the Kinect make it to be a primary device in 3D scene reconstruction, object recognition, image content recognition, and 3D points texture [18].

Skeletal Tracking allows Kinect to identify people and follow their movements. Within the field of view of the sensor, the infrared camera of Kinect can recognize maximum six users. However, it can only show tracking details of two users. Figure 2.8 shows the multiple users who are tracked by Kinect at the same time [19].

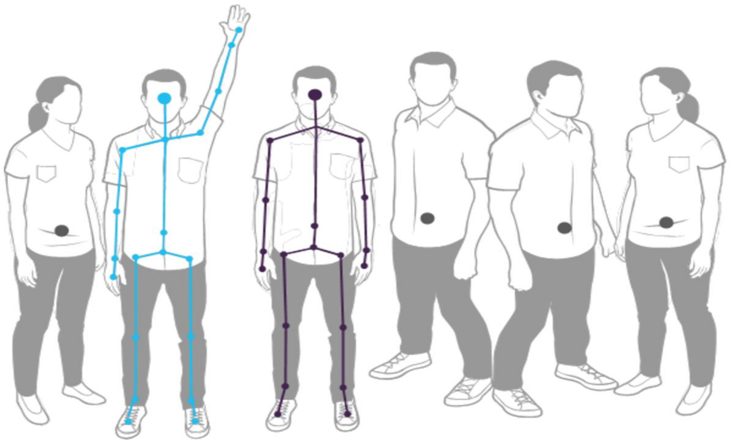
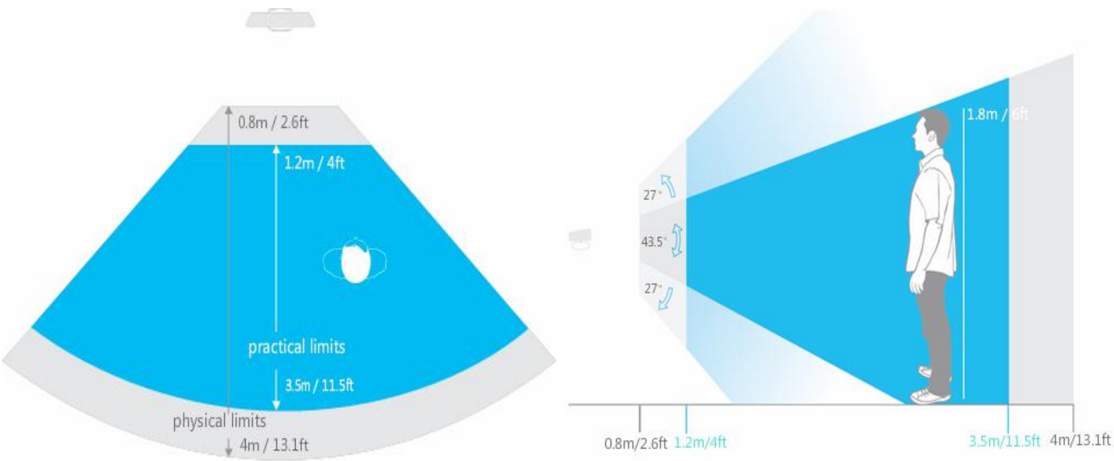


FIGURE 2.8: Six different users are recognized by Kinect [19]

User's field of view in Kinect is set on by the configuration of the IR camera. In the general default range mode, the possible distance between users and Kinect, which is between 0.8 and 4.0 meters, is designed to be recognized by the Kinect. Making the Kinect work accurately, the suggestion distance is from 1.2 to 3.5 meters. Figure 2.9 shows the horizontal and vertical fields of view in default range [19].

Starting from SDK 1.5, Kinect for Windows provides joint orientation data for the skeletons. The orientation data is offered in the form of quaternions and rotation matrices for utilizing in individual scenarios [20].



(A) Horizontal field of view in default range (B) Vertical field of view in default range

FIGURE 2.9: Kinect horizontal and vertical Field of View in default range [19]

A set of bones is clarified by the Kinect, using the joints defined by skeletal tracking system. The hip centre joint is the root and extends to the different joints, containing hand, feet and head. The Figure 2.10 shows the joint hierarchy in Kinect skeleton tracking system [20]. In this thesis project, each joint has an index for processing skeleton data in Unity.

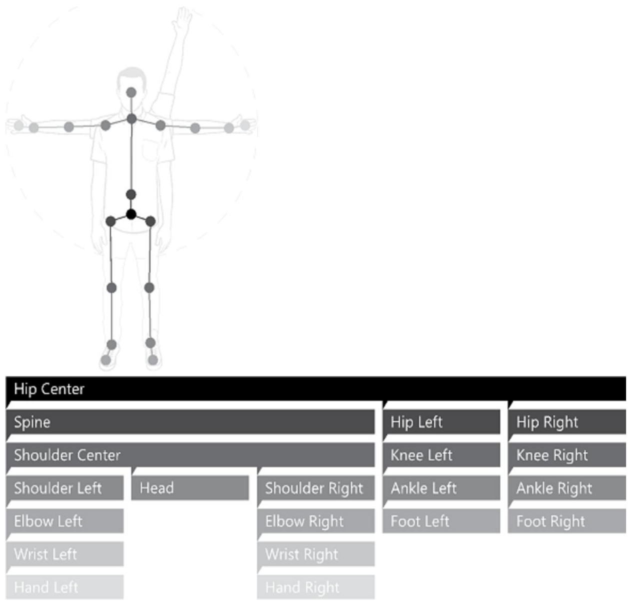


FIGURE 2.10: Joint Hierarchy [20]

Kinect is widely used by many researchers to collect the real-time skeletons tracking data. The Kinect in Augmented Mirror system performes user tracking in real-time, with an 11 bones avatar shape [21]. Kinect-based skeleton tracking is also utilized to evaluate the performance of dancer with the OpenNI drivers/SDK and is OpenNI-encoded [22]. Moreover, the Kinect is quite popular in interacting with users in game activities, with its significant performance in full body movements capture in 3D space.

Frati and Prattichizzo [23] worked in hand tracking and rendering in wearable haptics with Kinect, even before the Microsoft released the Kinect SDK. They chose the CLNUI platform with essential functions to retrieve data from hardware and to manage the Kinect's motor. They used the developed calibration procedure to calibrate parameters of Kinect data. The algorithm they used for hand tracking had been divided into four steps.

1. Process the depth image
2. Calculate the hand bounding box
3. Obtain main points of the hand including fingertip position
4. Filtering

The index and thumb tip measurements were provided by Kinect as well, as shown in Figure 2.11. Based on the data came from hand tracking algorithm, the avatar animation could be augmented to the hand. Their qualitative results presented the

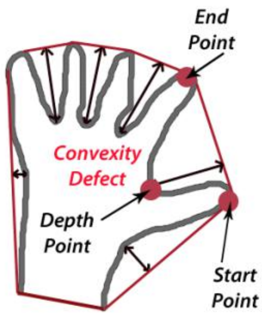


FIGURE 2.11: The Kinect processes the index and thumb tip measurements [23]

efficacy of combining Kinect and wearable haptics, while Kinect depth sensor offered the most important relevant data in the research.

To improve the AR experience, Clark et al. [24] explored the use of the Kinect, and it contributed in two main parts.

- Provide 3D information about the environment, which was integrated with the Kinect.

- A more realistic environmentally aware AR application was developed.

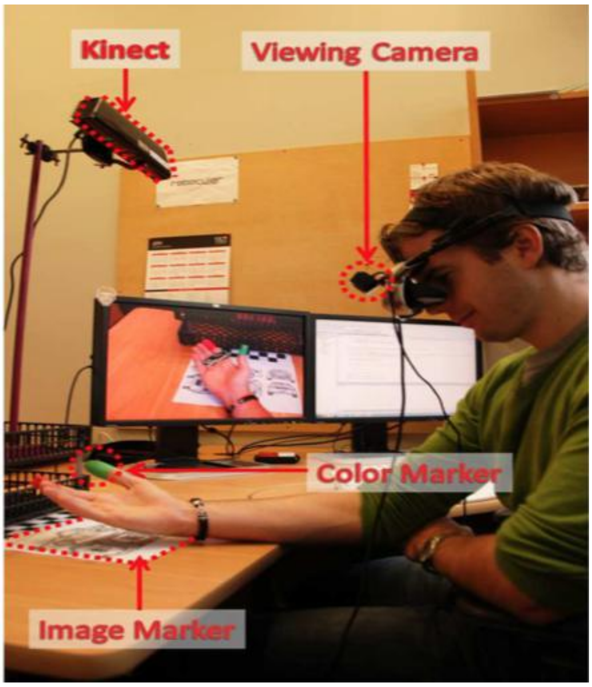


FIGURE 2.12: Setup of interaction [24]

Figure 2.12 shows the setup of Kinect and other devices. The user wore a viewing camera, and the Kinect was fixed above the interaction panel, while the image maker was below it. The Kinect transferred its 3D data to the AR co-ordinate system. Real-time updates of the environment were captured by Kinect, which enabled the environmental awareness and interaction methods with virtual content to be cheap and accessible [24].

The setup in Clark's research offered a working hardware setup example for this thesis project, which contained both Kinect and another camera SoftKinetic DS325.

2.4 3D Modelling

3D scanning has been a prevalent application of Kinect. An interactive reconstruction system was created, called KinectFusion [25][26]. Newcombe et al. [25] introduced two main features of the system: real-time surface mapping and dense real-time tracking. This system allowed the user to move within a specific space and reconstructed a scene with the high-quality 3D model, while the 3D model could be mapped with texture using the RGB camera of the Kinect. The team list several features of the KinectFusion, which are shown in Table 2.3 [26].

Table 2.3: Features of KinectFusion

Functions	**Features**
Scan	low-cost scanning be held in the hand
Segmentation	Object Segmentation via Direct Interaction
AR	Geometry-Aware Augmented Reality
Physics	Taking Physics Beyond the 'Surface'
Scene	Reaching into the Scene

The KinectFusion can be used as a low-cost scanner. The system allows users to capture an object from the different point of views quickly, even to reverse it, and show the feedback screen immediately. It also allows the user to segment specific objects from the entire scene by moving it physically. A more realistic way of AR which the physical world interacted and overlaid with the virtual world. Based on the simulated parts of the physics from real-world, dynamic interactions with the reconstructed scene are available for the virtual objects. The KinectFusion system is extended to a dynamic scene, and it contains the ability to interact with physics-enabled virtual objects for users [26].

In addition to developing a real-time 3D reconstruction system, Izadi et al. [26] also proposed the procedures of implementation by generic programming in the graphics processing unit (GPU). They developed an example AR application which used Kinect-Fusion system as a development tool(Figure 2.13) [27]. Figure 2.13A shows the scene which was captured, and the Figure 2.13B shows the 3D model of the scene, as well as the position of the camera. Several virtual particles could be added to the scene, and the environment can be reconstructed (Figure 2.13C). This immersive application also allowed interactions with the scene.

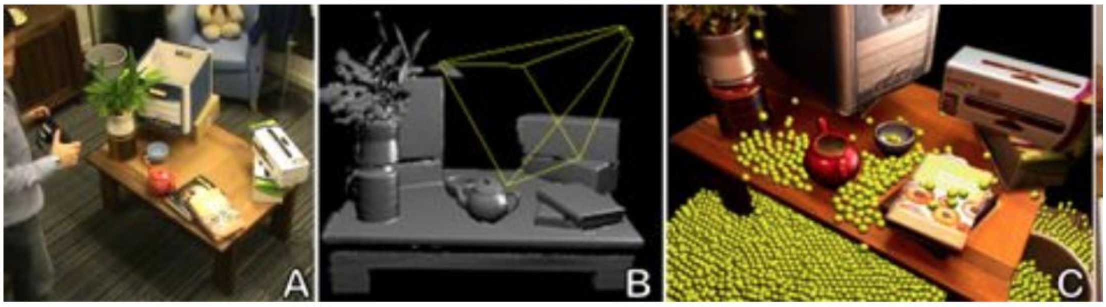


FIGURE 2.13: KinectFusion application [27]

These inspired the researcher to use the Kinect to generate the 3D construction and import into the Unity game engine.

There is another one to create a 3D avatar modelling. The word "Avatar" originally comes from ancient religious words, which has been used to describe as a controllable 3D embodiment of the user [28]. However, it has been widely used in the virtual world, representing a particular virtual character.

Aitpayev and Gaber [29] introduced their research of creating and animating a 3D Avatar by using Kinect with Microsoft SDK and OpenNI. They used a human predefined fitted 3D model to match the user's body in body creation. For Head (face) creation, they followed the below steps:

1. Face feature region detection
2. Face segmentation
3. Non-rigid registration
4. Deformed template mesh
5. Average face overlaid
6. Final reconstruction and texture

Furthermore, there were two parts to animate the created Avatar as well [29]. Brekel Kinect 3D program, which allowed to capture 3D objects and output them to the disk for 3D packages, offered a most straightforward way to complete the body movement animation. Predefined animation and real-time performance-based animation were both needed in facial animation.

Weise, Bouaziz, Li and Pauly [30] had presented a new algorithm for facial animation. It acquired the 2D image and 3D depth map from Kinect. A tracking algorithm, applied by user-specific expression model, probabilistic animation prior and temporal coherence, had been used to adjust the Blendshape Weights before completing the animation. Figure 2.14 shows the processing pipeline.

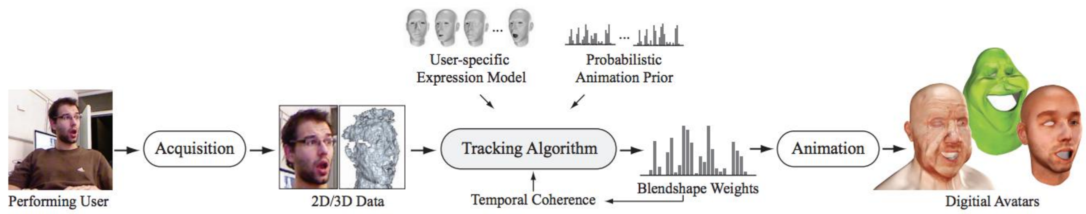


FIGURE 2.14: A process pipeline of real time performance-Based facial animation [30]

Augmented Mirror, an interactive AR system based on Kinect [21], was used to interact a virtual actor with an audience. It contained control scenario and augmented scenario, using a client-server model to connect both. An enhanced MoCap system was created to collect input data and processed via a server application. The Avatar movement was controlled by the main device of MoCap system which was Kinect depth sensor camera with Open Natural Interface (OpenNI) SDK. However, the real-time tracking was not accurate enough, four ways of enhancement had been introduced:

- head orientation

  Attached mobile phone to the cap was used to analyze the head orientation. Based on the user's orientation, mobile phone's sensor data, global reference data including earth gravity and the North Pole, an Android application could be used to calculate the head orientation.

- lips movement while talking

  A wireless microphone was given to collect and compute the horizontal and vertical number of the lip according to the simulation of lip movement. In this case, an amplitude based algorithm had been created.

- a WiiMote with control algorithms facial expressions

  There were five different predefined facial expressions set in the system initially.

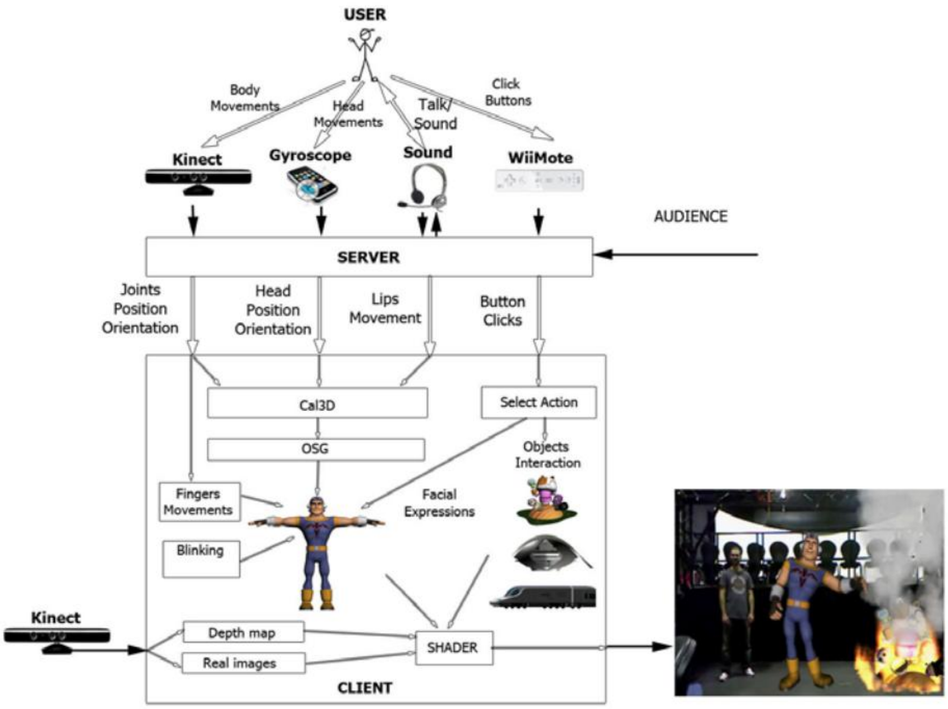


FIGURE 2.15: The overall structure of the Augmented Mirror system [21]

- automatic gestures

  These gestures including, but not limited to hands, feet, blink, etc. All the data would be calculated automatically and combined with other data later.

All the collected data provided the essential fundamental for real-time augmentation. The mixed real image and depth data captured by Kinect were merged on a GPU shader with virtual scene and elements to implement the augmented mirror image. Customized libraries, Cal3D and OpenSceneGraph, were all used to achieve the augmented visualization. Figure 2.15 describes the overall structure of the application.

There is another way to transform the Kinect image data to the Unity described in previous work by Rudhru [31] A plugin was developed using Visual Studio C++ 2010 with Microsoft Kinect SDK v1.7 in the prototype. The Kinect would provide 1280*720 resolution of RGB image system and 640*480 resolutions of depth image stream. This plugin allowed the Unity 3D to access the Kinect sensor data which is provided by the Kinect SDK. The software architecture is shown in Figure 2.16.

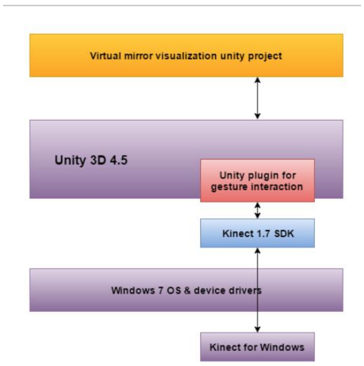


FIGURE 2.16: The software architecture of Rudhru's prototype [31]

2.5 Game Engine

A game engine is a collection of interacting software [32], in a single unit runs an actual game, consists of several subsystems with specified functionality. The main subsystems contains audio, input, physics, rendering, artificial intelligence (AI), core, scripting, and networking (Figure 2.17).

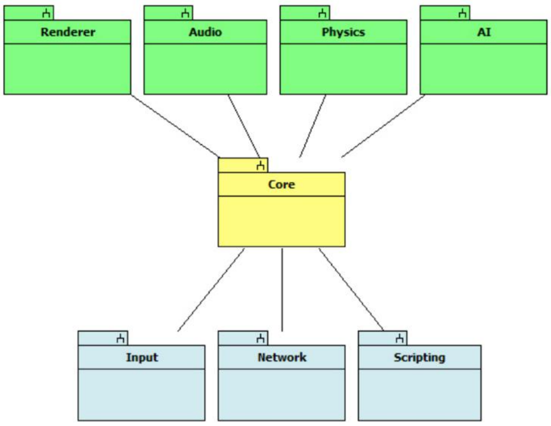
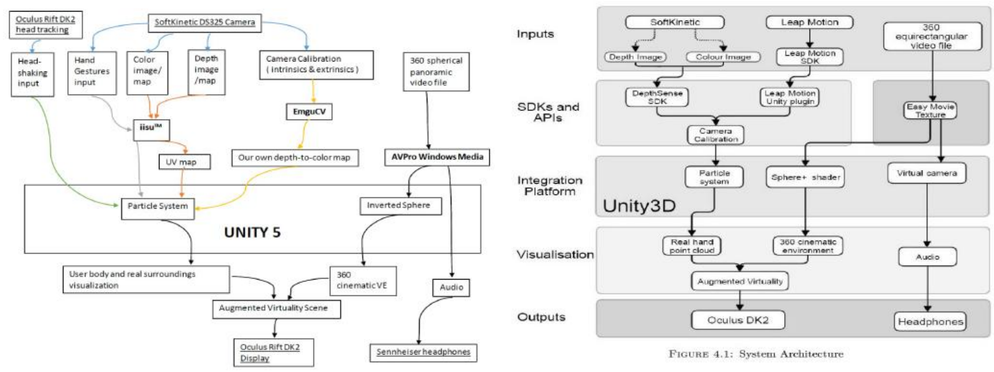


FIGURE 2.17: The abstract design of the game engine [32]

Shiratuddin and Thabet[33] created a virtual office walkthrough using a 3D game engine in 2002, which pointed out that 3D game engine offered a low-cost VR solution because of its significant built-in features, such as rendering, collision detection, sound, scripting, and animation. There were games using Unity game engine to create an immersive virtual environment. Bourke [34] discussed an immersive display environment in a hemispherical dome and presented a method of creating the correct projections with Unity game engine. Jorge and Couto [35] discussed using the Unreal game engine in evacuation planning. Benko et al. [36] also used Nvidia PhysX physics engine to do their physics simulation for their MirageTable project, which provided freehand interaction on a projected AR tabletop.

There are a large number of games and projects using the game engine. Independent scenarios and open source codes offer good support for researchers or developers to do VR simulation using game engine [35]. In previous work, Unity [37] game engine has also been used as essential parts of the system. The latest version of Unity game engine has the built-in support for VR HMDs. It becomes convenient to integrate essential devices together.

Unity was used as the principal software platform to process the interaction of data and HMD in Chen's [5] and Khan's [38] previous works, as well as the work done by Rudhru [31]. Therefore, Unity 5 is the leading software tool for this thesis project. Figure 2.18 shows the system architectures from Chen's and Khan's prototype.



(A) System architecture of Chen's prototype [5]

(B) System architecture of Khan's prototype [38]

FIGURE 2.18: System architectures from previous work

2.6 Summary

This chapter described the previous related work to this thesis. It contained the popular concepts, projects, and algorithms in MR cinematic experience, AV, Microsoft Kinect, Image composite and modelling, and Unity game engine. Next chapter will describe the details ideas and design procedure of this thesis study.

Chapter 3

Design

This chapter describes the idea formation of adding AV experience to the immersive cinematic experience, as well as the plan of achieving this goal with one of HMDs named Oculus Rift DK2, two depth cameras which are Kinect and SoftKinetic DepthSense 325.

Recent technology development in HMDs is making immersive VR experiences easily affordable for the general public. It brings immersive cinematic experiences from the public 3D theatres into home entertainment (Figure 3.1).

FIGURE 3.1: Immersive cinematic experience

3.1 Design Concept

The overall goal is to take the personal immersive cinematic experience further to the next level, where the users can have more intimate experiences by perceiving themselves as being part of the movie scene. The general design idea is to improve prior work by Chen 2017 [5] to achieve the proposed goal of augmenting the virtual objects (including user's body) in the AV with matched themes of the cinematic experiences. Figure 3.2 shows the concept design of augmenting the user's body into the cinematic scene, while Figure 3.3 shows the concept design of a system through which users could experience immersive 360° movies while seeing their own body blended into the scene and interacting with virtual objects embedded in the movie scene.

FIGURE 3.2: System concept scenario

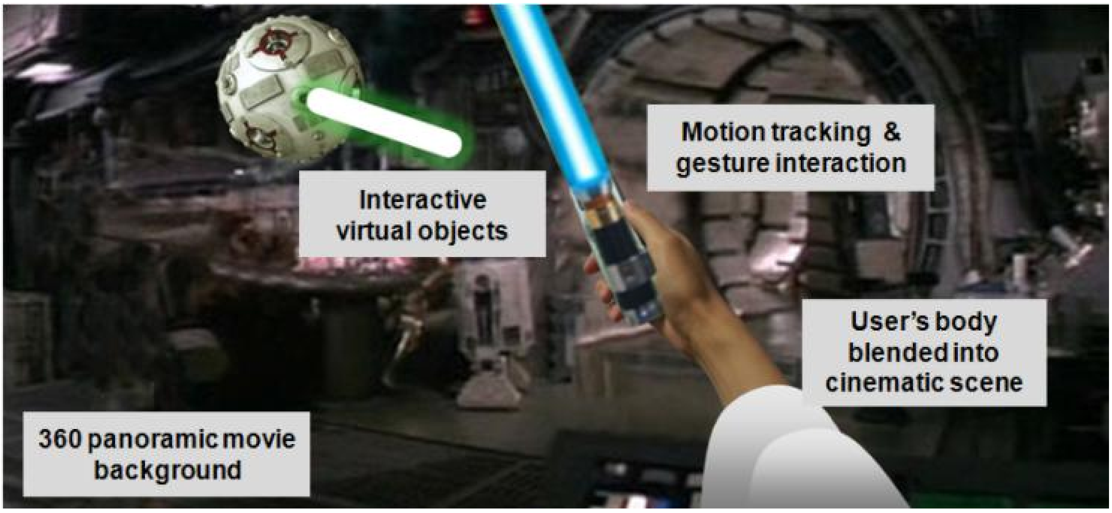


FIGURE 3.3: System concept scenario

3.2 Design Consideration

There are a variety of animations that can be used to enhance users' immersive cinematic experiences, including but not limited to weapon, mask, armor, tattoos, clothes.

Meanwhile, these animations can be augmented to different parts of the real body, such as face, arms, hands, legs, feet and body.

Initially, the proposed system consisted of getting Unity to communicate with the Kinect depth camera, receiving depth image and Kinect Fusion data, displaying in the Unity. Kinect Fusion enables the generation of real-time detailed 3D reconstruction scenes by holding and moving the Kinect camera. The Kinect camera can capture depth data from multiple viewpoints. The Kinect Fusion system will integrate these data to create a smooth single dense surface model. Figure 3.4 shows the critical steps of processing pipeline from raw depth to a 3D reconstruction [39].

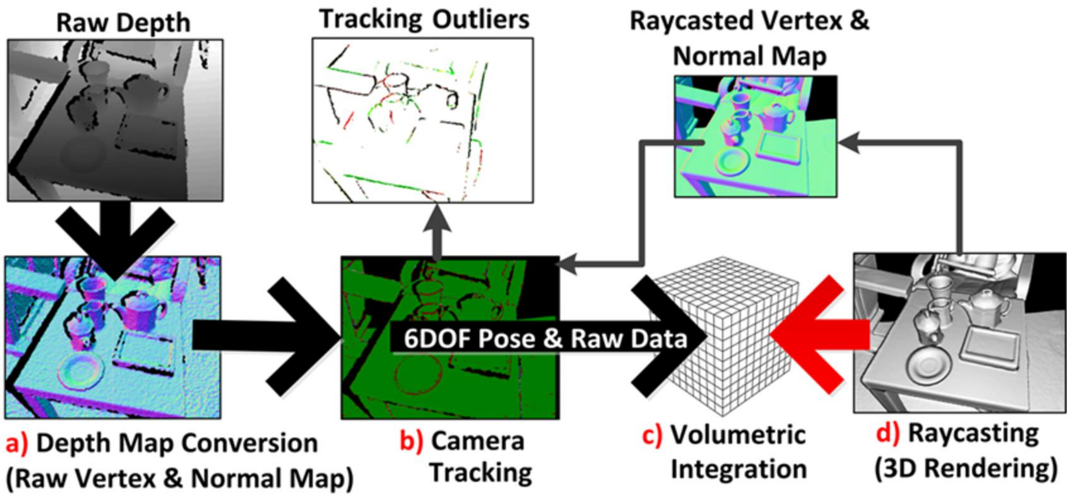


FIGURE 3.4: Processing pipeline of KinectFusion

To reconstruct the physical environment near the user, the environmental depth-sensing camera has been used. Figure 3.5 shows the colour and depth images captured using the Microsoft Kinect depth-sensing camera facing the user wearing a HMD that was in the prototype system setup from Chen's [5] prototype.

FIGURE 3.5: Colour (left) and depth (right) image captured from Kinect[40]

Kinect SDK is used to run Kinect Fusion 3D reconstruction algorithm, which uses a stream of depth images to form a volumetric representation of the reconstructed surface. Figure 3.6 shows the final mesh representation of the reconstructed physical environment around the user.

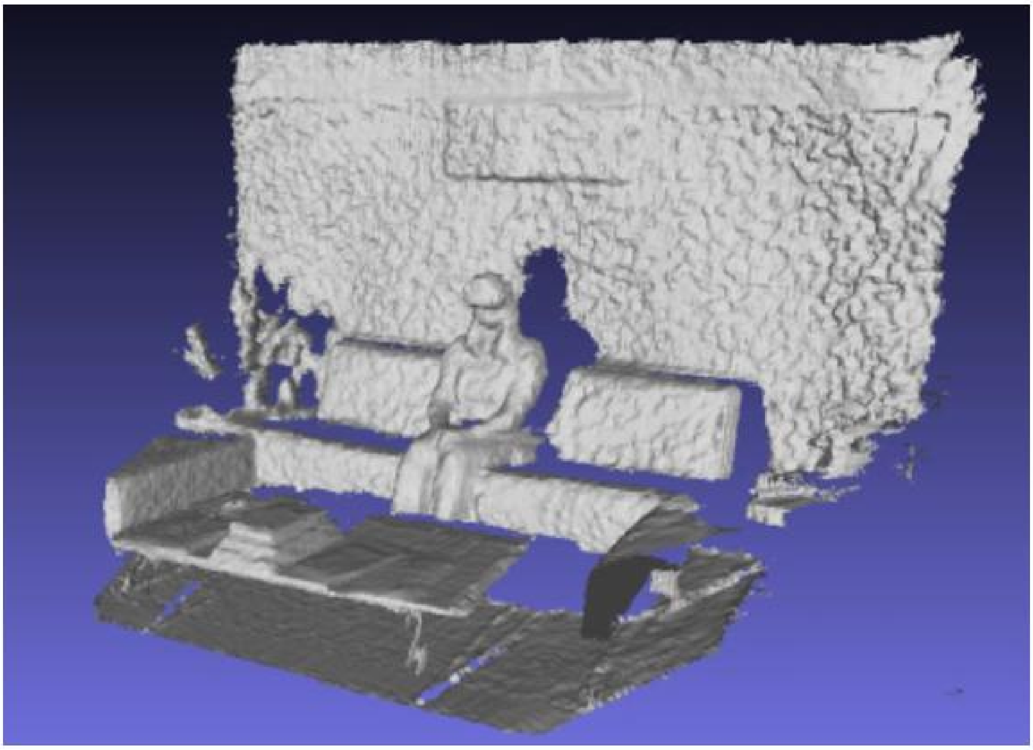


FIGURE 3.6: KinectFusion 3D data captured from Kinect

However, after integrating the reconstruction module into a Unity project, the researcher decided not to continue using Kinect Fusion because of the following reasons:

- The user had to stand and move the Kinect when reconstructing the surroundings, which could not create a consistent experience when watching the movie.

- The speed of reconstructing, the quality of the image and the stability of transforming reduced the efficiency of the proposed system.

- The design and modelling skills of 3D rendering were more complicated than expected which might slow down the progress of the project.

In this thesis, the Kinect is predominantly used to collect user's skeleton tracking data, while the SoftKinetic camera was principally utilized to render the real body data. Therefore, the designed system is composed of two main depth sense cameras, the Kinect and the SoftKinetic DepthSense 325.

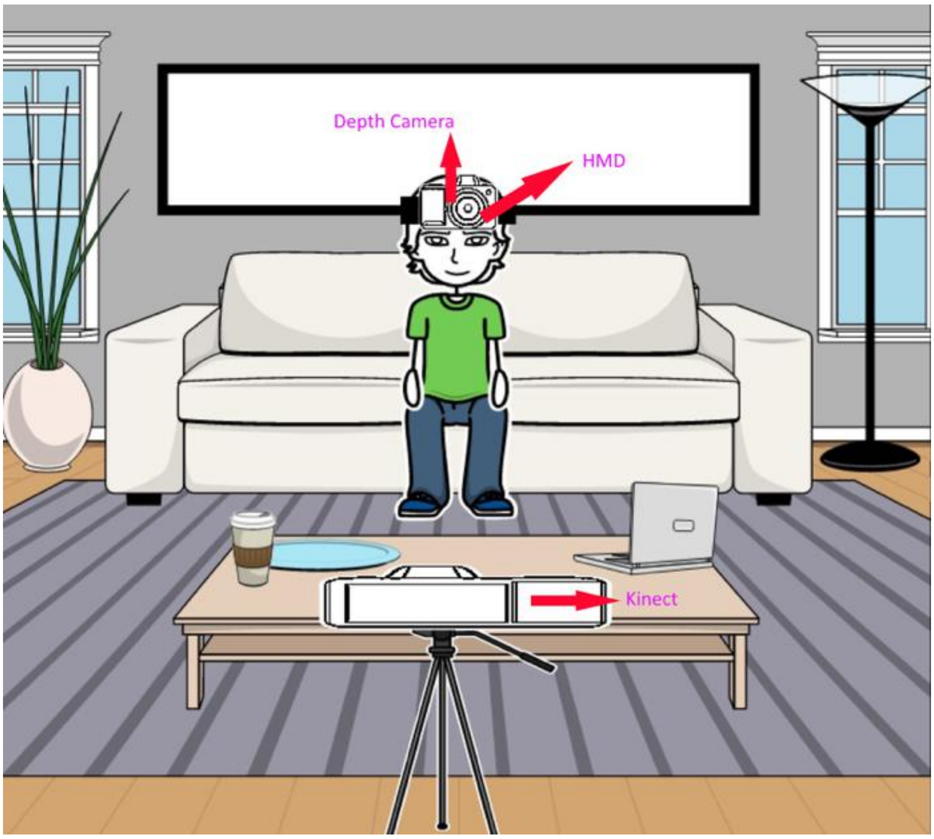


FIGURE 3.7: Hardware setup sketch

According to previous related work, the Kinect presents a high performance in the user's skeletons tracking. Hence, the part of the body that has more tracking joints will be prioritized in the augmented scenario.

The study focuses on immersive cinematic experience, which means the augmented objects should be related to the 360° movies. Therefore, the content-based objects created are required to be augmented on a real body. Existing 3D models could be acquired on an open source website or database to match the video content.

3.3 Prototype System Setup Sketch

3.3.1 Hardware setup sketch

To show the improvement of visualization of the physical environment, a system prototype has been developed. The prototype sketch aims to capture the whole environment around the user and the user himself when watching the 360° video, as well as rendering physical objects which would be augmented based on the theme of the movie in real-time. In this system, users can see themselves and the physical objects in a more natural way of blending into the cinematic scene.

The primary interfaces used in the immersive MR cinematic experience prototype system contains a Kinect, an Oculus Rift DK2, and a SoftKinetic DepthSense 325. The SoftKinetic camera will be mounted on the Oculus Rift DK2 while the Kinect will be put on the other side. Figure 3.7 shows the sketch of how the devices are set up.

The Kinect is used to capture the broader view around the user, including the user himself and his skeleton, as well as the surrounding physical environment. The depth camera attached to the HMD will capture the physical objects from user's perspective.

3.3.2 System Architecture Design

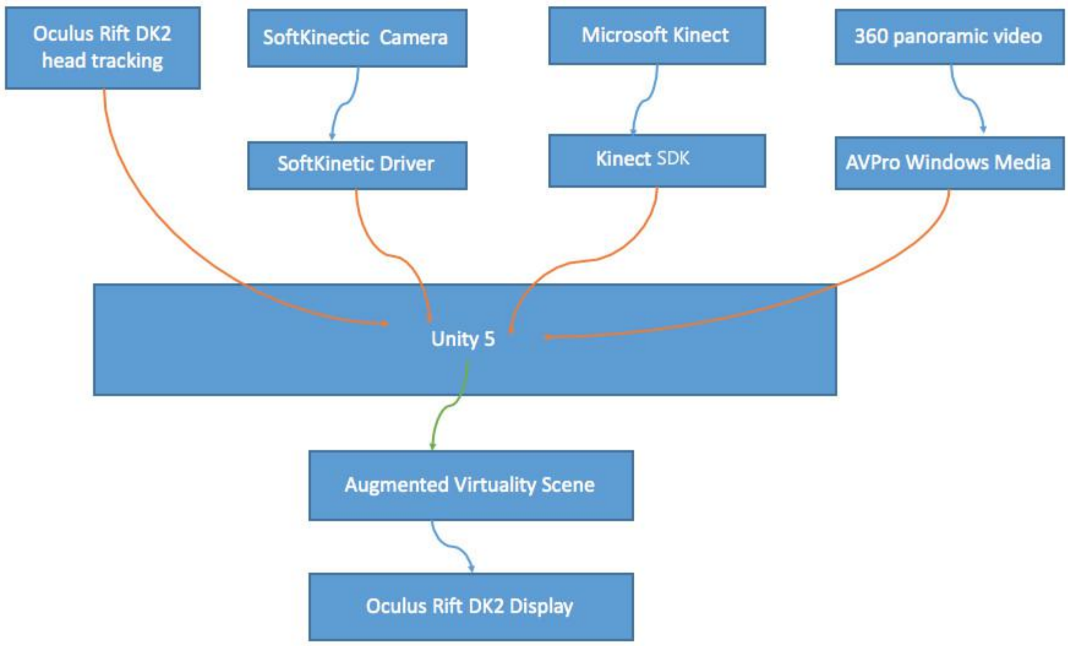


FIGURE 3.8: System Architecture Design

The virtual environment is mostly implemented in the Unity 5 game engine. The Unity 5 game engine has a built-in feature supporting the Oculus Rift DK2 which eases the software design. The SDK of the Softkinetic camera and the Microsoft Kinect need to be installed and well explored.

Realistic or natural textures need to be created or found, and specific 360° movies need to be downloaded and imported into the Unity 5 game engine. Figure 3.8 shows the software modules and tools which would be used in the project.

3.4 User Experience Design

The user experience is the critical part of the prototype. There are several scenarios which the researcher has considered.

The first one is providing several different types of 360° movies. These different movies will be augmented with different virtual objects. The advantage of this is that the user will be more interested because of more abundant contents available. The disadvantage is that it will be difficult to do the pair-comparison.

The second one is using the same 360° movies with different virtual objects augmented. Similar to the first scenario, it would be good for attracting the user's attention, but hard to conduct a pair-comparison.

The third one is using the same 360° movies with the same blended virtual objects. It may distract the user's attention but will allow an easy pair-comparison.

FIGURE 3.9: Example for visual effects and adjustments [41]

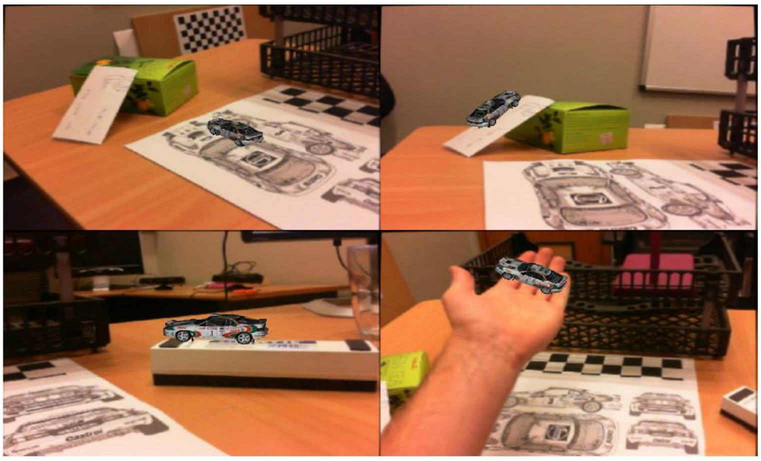


FIGURE 3.10: Example for visual effects and adjustments [41]

There are different methods that can enhance the visualization as well. One is adjusting the visual appearance of the video overlay, including hue, saturation, brightness and transparency values of the corresponding pixels (Figure 3.9) [41]. Another one is augmenting some virtual objects on top of the real objects, as shown in Figure 3.10 [24].

3.5 Summary

This chapter mainly outlined how the idea of this study had been generated, as well as the initial design of the working system. The hardware setup , software structure and user experience design have been described. In the next chapter, the implementation of the prototype system will be explained.

Chapter 4

Prototype Implementation

This chapter explains the implementation of the prototype which the researcher used in the study. Following the initial design consideration and the system sketches, the working prototype was built which mainly consists of two parts, hardware, and software, combining two depth cameras (Kinect and SoftKinetic DepthSense 325), a HMD (Oculus Rift DK2), and a headphone with 360° VR movies in Unity 3D environment.

4.1 System Architecture

Figure 4.1 presents the system architecture of the study prototype. This block diagram demonstrates the comprehensive procedures and architecture of the working prototype.

There are four layers to the working prototype system, which are input, SDK and API, integration and visualization, and output.

The system outlines four input source components. Oculus Rift DK2 Head sensor provides the head orientation, which helps the systems to define the skeleton position. With the help of Kinect for Windows SDK v1.8 and the AM plugin script, the skeleton data is transformed to the Unity 3D. Additionally, the SoftKinetic camera can deliver real body image via its depth sensor SDK. 360° videos are passed into the Unity 3D as well.

It also describes the output AV source in Oculus Rift DK2 and the headphones following the integration and visualization operations in Unity 3D.

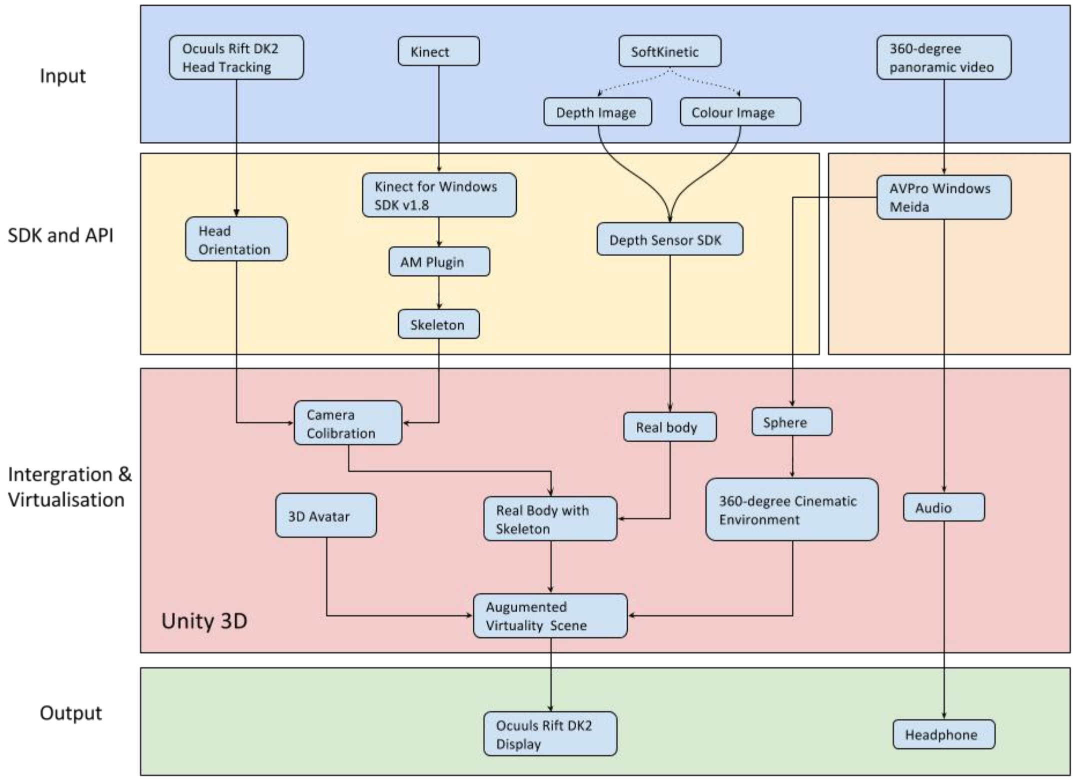


FIGURE 4.1: System Architecture

4.2 The Hardware Requirements

The hardware requirements for this thesis is similar to prior work [5]. The Oculus Rift HMD is used, on which is mounted a SoftKinetic RGBD camera with a 3D printed frame. One more depth camera is needed to track skeleton information for this project, which is the Microsoft Kinect. General hardware also contains a computer with the game engine installed. A pair of proper headphones for sound effects of the movie is also prepared.

4.2.1 Kinect

The Kinect sensor is designed as a black box which places on a small platform, ideal to put down a desk. The specifications for the Kinect are shown in Table 4.1.

Essential native and managed tools and APIs are provided by the Kinect for Windows SDK, which the user could develop Kinect-enabled applications for Microsoft Windows. In this project, Kinect for Windows SDK 1.8 is chosen to be used.

TABLE 4.1: Specifications for the Kinect [42]

Kinect	Array Specifications
Viewing angle	43° vertical by 57° horizontal field of view
Vertical tilt range	±27°
Frame rate (depth and color stream)	30 frames per second (FPS)
Audio format	16-kHz, 24-bit mono pulse code modulation (PCM)
Audio input characteristics	A four-microphone array with 24-bit analog-to-digital converter (ADC) and Kinect-resident signal processing including acoustic echo cancellation and noise suppression
Accelerometer characteristics	A 2G/4G/8G accelerometer configured for the 2G range, with a 1° accuracy upper limit

4.2.2 SoftKinetic DepthSense Camera

SoftKinetic DepthSense 325 camera is the end-to-end provider of natural gesture recognition solutions, delivered real-time 3D distance data for close interaction. It provides depth data for software analysis from as close as 15cm, at up to 60fps. It contains a DepthSense sensor, an high definition RGB sensor and two microphones.

Figure 4.2 shows the technical description of the SoftKinetic DepthSense 325 camera , while Figure 4.3 shows the product specifications [43].

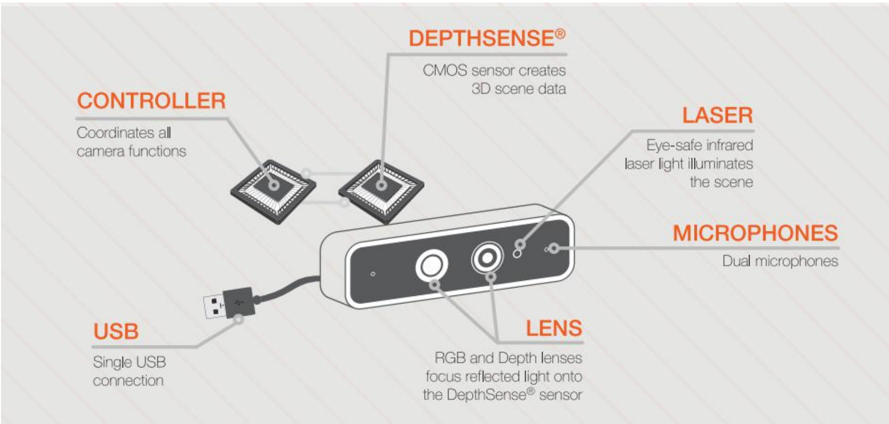


FIGURE 4.2: SoftKinetic technical description [43]

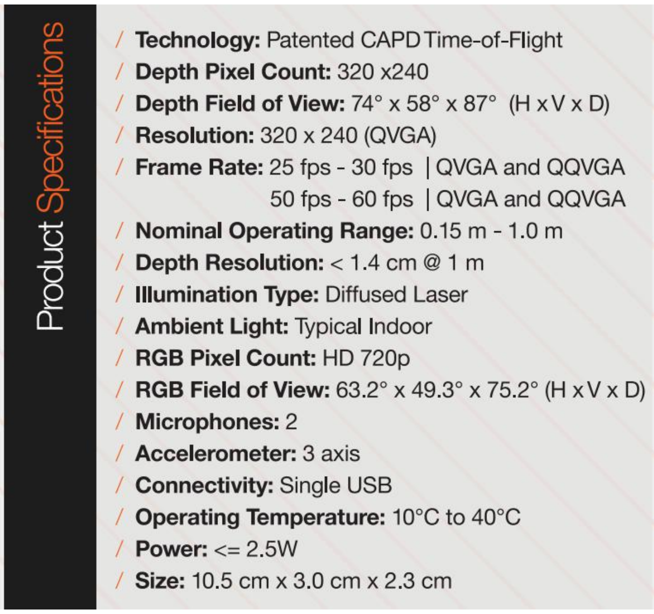


FIGURE 4.3: SoftKinetic product specifications [43]

4.2.3 The Head-mounted Display

In the project, Oculus Rift Development Kit2 (DK2) is used as a HMD that allows developers to build various games and experiences [44]. There are two components which have been used in this project, a headset for display and a sensor for tracking (Figure 4.4).

(A) The headset (B) The sensor

FIGURE 4.4: Two components of Oculus Rift DK2 used in the study

The Oculus Rift DK2 has recommended computer specifications requirements to power it, as shown in Table 4.2.

Table 4.2: Recommended Computer Specifications for Oculus Rift DK2 [45]

Recommended Specifications	
Graphics Card	NVIDIA GTX 1060 / AMD Radeon RX 480 or greater
Alternative Graphics Card	NVIDIA GTX 970 / AMD Radeon R9 290 or greater
CPU	Intel i5-4590 equivalent or greater)
Memory	8GB+ RAM
Video Output	Compatible HDMI 1.3 video output
USB Ports	3x USB 3.0 ports plus 1x USB 2.0 port
OS	Windows 7 SP1 64-bit or newer

4.2.4 Personal Computer

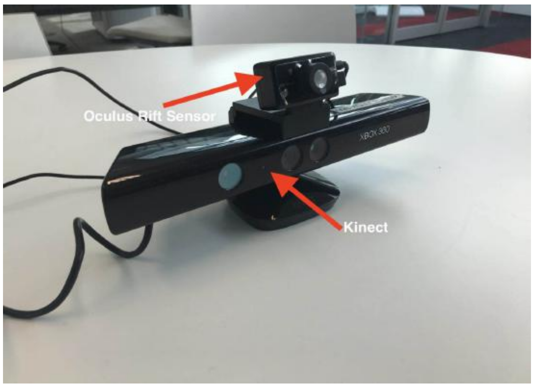
These recommended specifications for the Oculus Rift DK2 also matches the system requirements of Kinect and SoftKinetic camera. Therefore, the specifications of the computer used in this thesis project are shown below in Table 4.3.

Table 4.3: Computer specifications in this thesis project

Recommended Specifications	
Graphics Card	NVIDIA GeForce GTX970M + Intel GMA HD 530
CPU	Intel core i7-6700HQ
Memory	8GB+ RAM
Video Output	Compatible HDMI 1.3 video output
USB Ports	3x USB 3.0 ports plus 1x e-SATA 3.0 port
OS	Windows 7 SP1 64 bit

4.2.5 The Hardware Setup

(A) Oculus Rift with SoftKinetic camera



(B) Kinect with Oculus sensor

Figure 4.5: Hardware setup for the prototype

As described in previous subsections, the SoftKinetic camera would be mounted on the Oculus Rift DK 2, and the Oculus sensor would sit on top of the Kinect. These are shown in Figure 4.5.

4.3 The Software

The software in this study contained 4 parts. Unity 3D game engine, Kinect for Windows SDK 1.8, Depthsense SDK, and SketchUp for 3D models.

4.3.1 Unity Game Engine

The virtual environment is mostly implemented in the Unity game engine. Unity is a cross-platform game engine developing engaging 2D, 3D, VR, and AR apps and games. It enables users to perform fast prototyping and deploy the content to virtually multiple devices or media channels. It supports various platform. Therefore, it is convenient to connect Oculus Rift Dk2 to the Unity with built-in features [46].

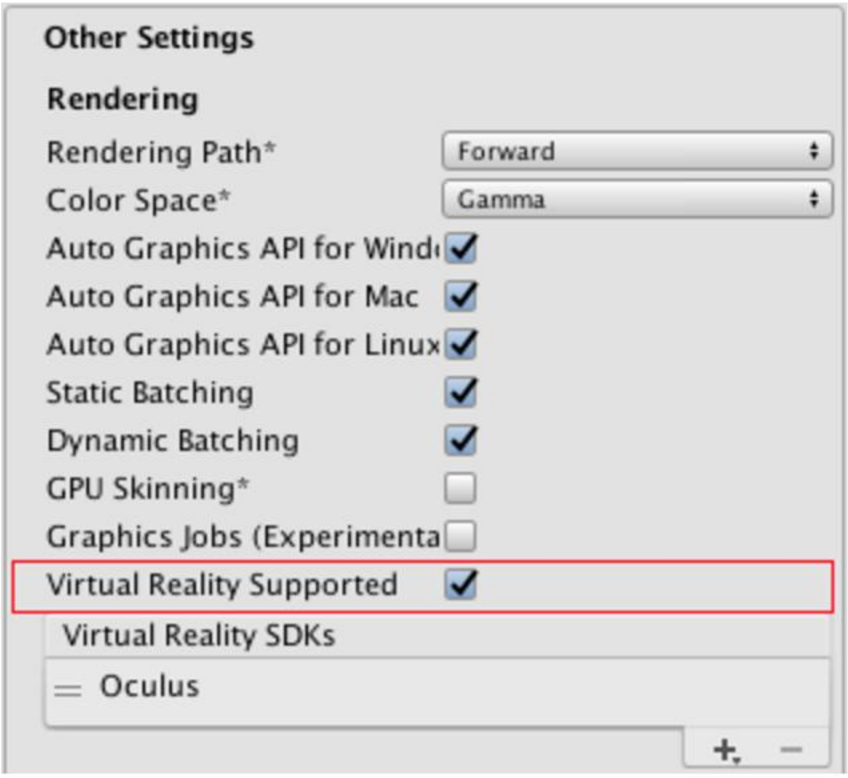


Figure 4.6: Other setting interface in Unity 5

The Unity is frequently updated, the latest version being Unity 2017.3, while the Unity version that was used for the project was version 5.4.6. This version was the most suitable version at that time when the study was done.

Unity VR makes VR devices easily connected directly from Unity, without installing any external plug-in in the projects. A primary Application Programming Interface (API) and multiple devices' compatibility of feature set are provided by Unity VR. It also designs for the future devices and software. The user can gain many benefits from native Unity VR support [47]. The native VR support feature can be activated by changing the Unity application setting.

When the VR setting is enabled in Unity, it automatically renders a HMD and head-track input, as well as understand the camera [47].

4.3.2 Kinect for Windows SDK 1.8

Kinect for Windows SDK 1.8 is an essential SDK to use if the Kinect-enabled applications are developed in Microsoft Windows. It allocates both native and managed APIs and tools. Figure 4.7 shows the interface of Kinect for Windows developer toolkit v1.8.0.

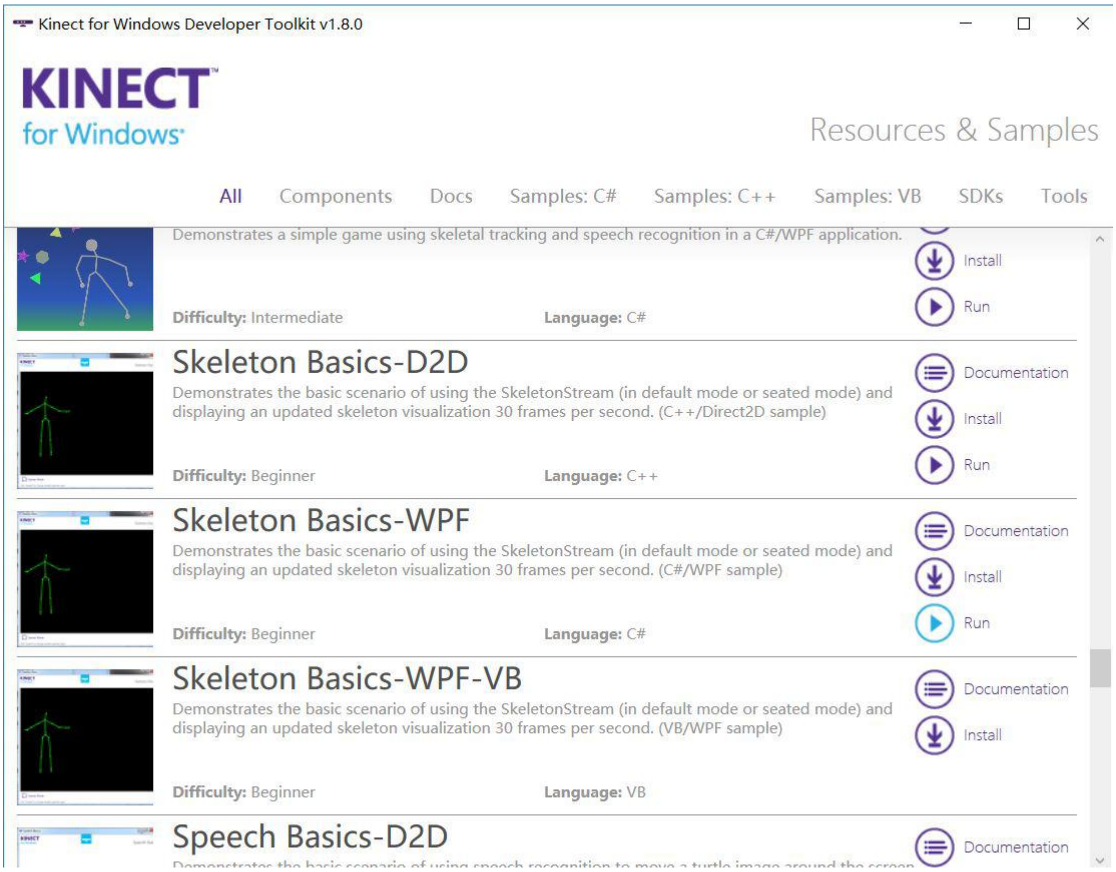
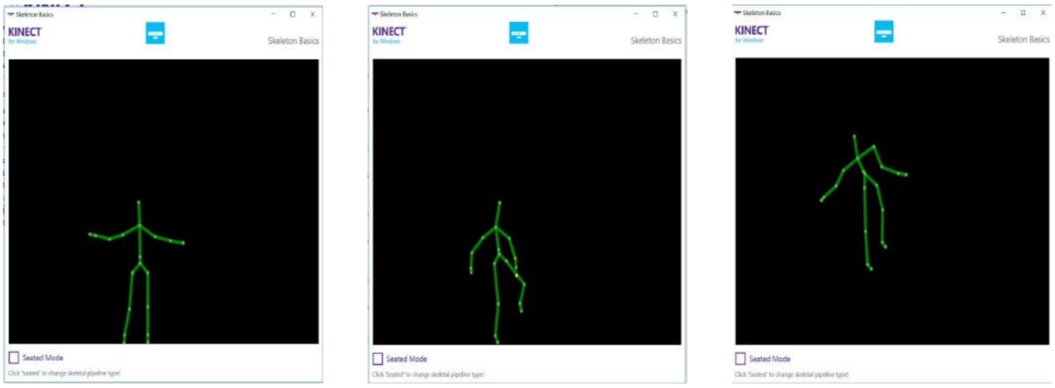


FIGURE 4.7: Kinect for windows developer toolkit v1.8.0 Interface

To test the Skeleton tracking feature, a "Skeleton Basic" sample has been selected. The joints of the user are shown on the screen in Figure 4.8. It shows the user's joints in three different positions. The sample application called the functions of the Kinect to collect

20 joints data information of user's skeleton tracking. These skeleton tracking data can be transferred to Unity via amended AM plugin script, which will be elaborated in a later section.



(A) Position 1 (B) Position 2 (C) Position 3

FIGURE 4.8: Showing user's skeleton of Skeleton Basis Sample in Kinect for windows SDK v1.8

4.3.3 SoftKinetic DepthSense SDK

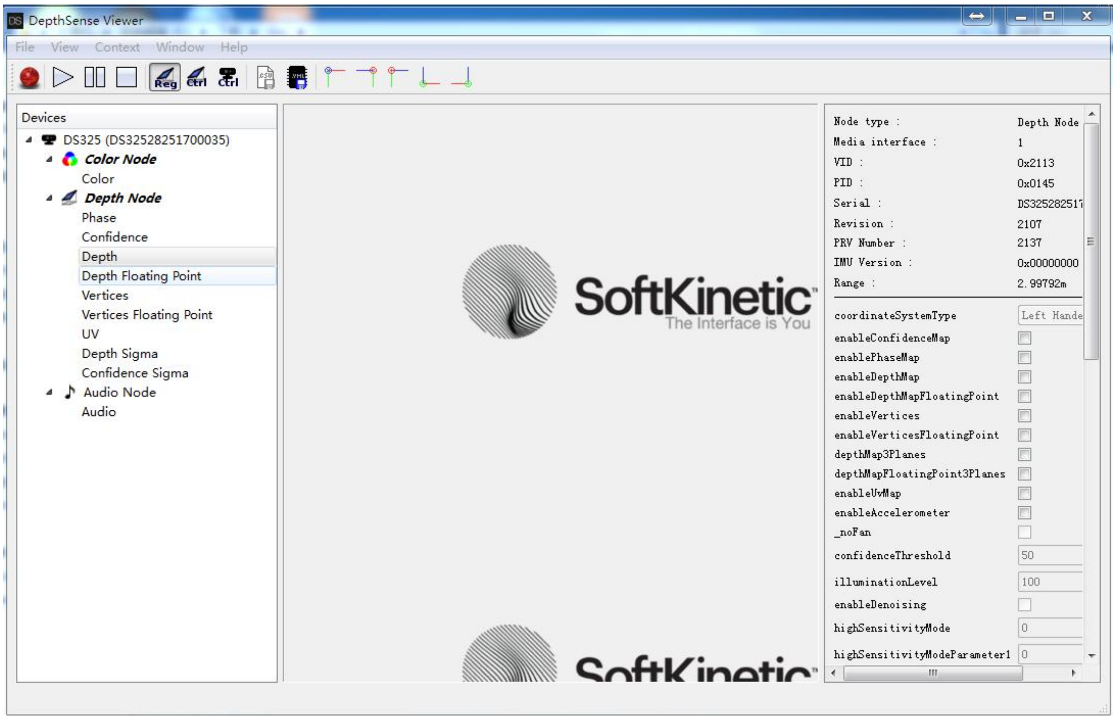
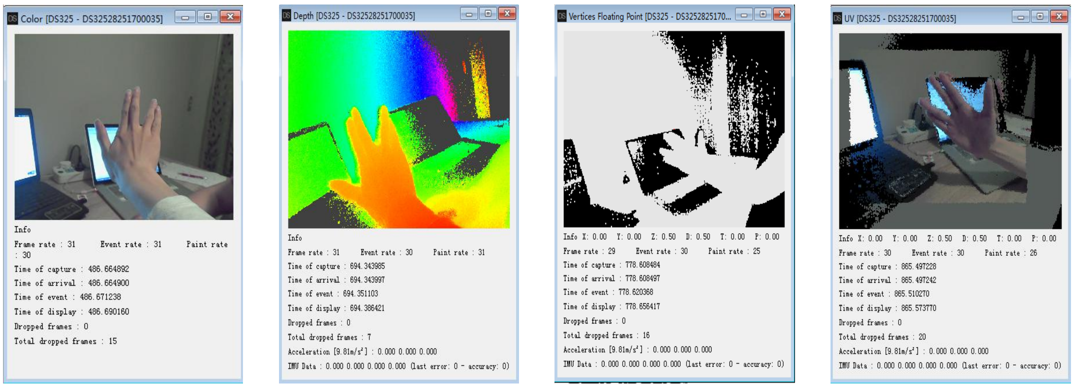


FIGURE 4.9: DepthSense Viewer Interface

The SoftKinetic DepthSense 325 camera is used in the study, with the help of DepthSense SDK, a user can access complete data from this camera, The viewer interface allows the user to configure settings in colour node or depth node (Figure 4.9). The most relevant setting is the range of the field of the camera, which decided how many real entities would be rendered into the virtual environment. Figure 4.10 shows how the RGB colour data, depth map, vertices map and the UV map are retrieved. To transfer compulsory data streams into Unity 3D, a hand gesture plugin has been used, which will be described in a later section.



(A) RGB colour data (B) Depth map (C) Vertices map (D) UV map

FIGURE 4.10: Camera data shown in Depth Sense Viewer

4.3.4 SketchUp

SketchUp [1] is a 3D modelling software. In SketchUp users can create 3D models, import and export the 3D models and share them with others. The 3D warehouse contains a lot of existing 3D models which are created by other users in the SketchUp. It is convenient to find some established 3D models for the project and customize them into pieces.

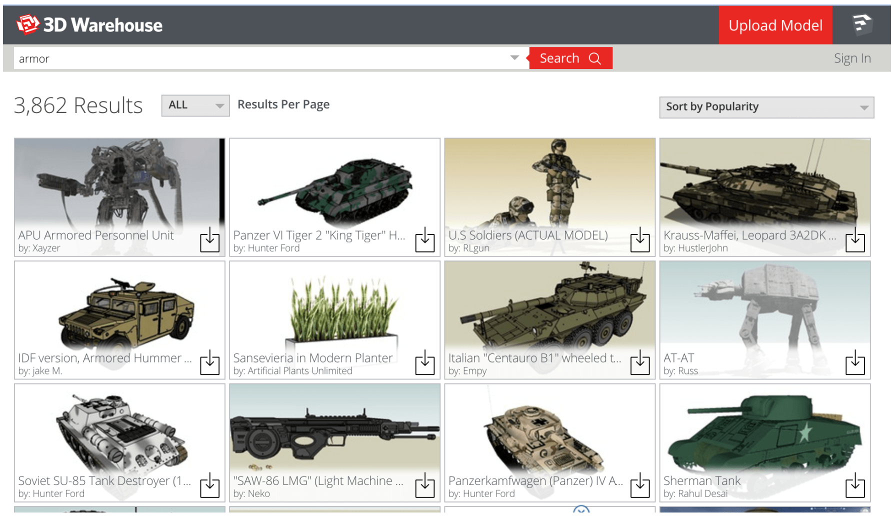


FIGURE 4.11: SketchUp 3D warehouse search result interface

Figure 4.11 shows the search result interface in SketchUp 3D warehouse. In this study, some 3D models of armors and weapons are imported to Unity which is explained in a

[1] https://www.sketchup.com/

later section. Figure 4.12 shows the interface in SketchUp Make which can generate and edit the 3D models.

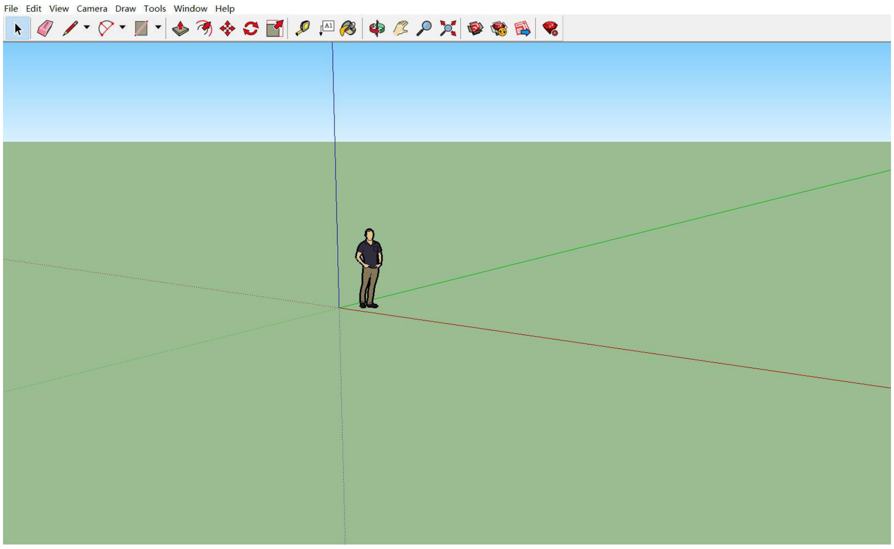


FIGURE 4.12: SketchUp 3D make interface

4.4 Prototype Integration

4.4.1 360° panoramic movie player in Unity 3D

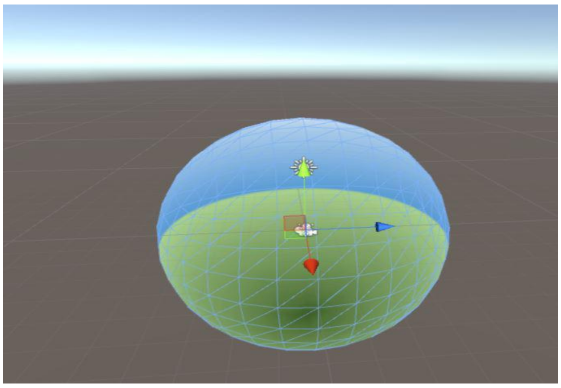
The 360° movie can be downloaded from external resources, including broad themes. It is generated by a collection of cameras systems that can captured all 360° of a scene simultaneously.

Either a 360° camera or a collection of cameras with overlapping fields of view can be operated to capture whole spherical field images. The specialized software tools are implemented to sew these captured images into a seamless panoramic video, which renders a 360° projection. The commonly used projections are equirectangular and cubical.

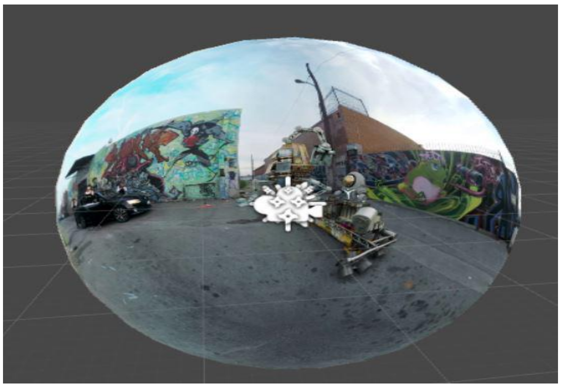
To activate the internal video player support feature, a video player component attached to a spherical GameObject needs to be generated. A "PanoramaVideoPlayerBehaviour" script is implemented to call the video clip asset plugin, to import 360° movie into Unity 3D. Figure 4.13 shows the 360° movies player mapping.

Table 4.4 compares the advantages and disadvantages of the different video player assets for Unity. Therefore, the video player asset the research used in the project is the "easy movie texture", considering the cost and the video viewing experience.

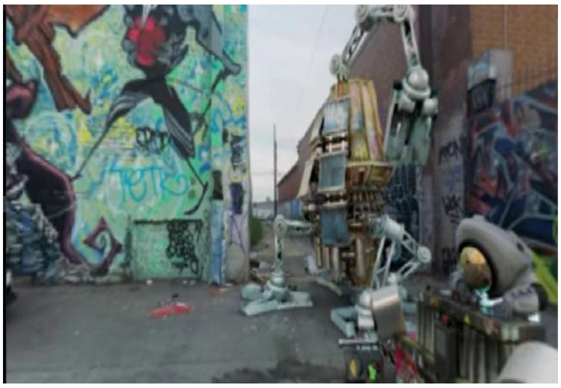
(A) Equirectangular projections



(B) Sphere generated in Unity 3D for mapping



(C) Video player with inside mapping on sphere



(D) Main Camera's perspective

FIGURE 4.13: The procedures of mapping 360° panoramic movie in Unity 3D

TABLE 4.4: The comparison of different video player assets

Assets	**Advantages**	**disadvantages**
AVPro Trial Version	Free	Has watermarks
Easy Movie Texture	Free	Reduce video quality
AVPro Full version	No watermark	Expensive

4.4.2 AM Plugin, Skeleton in the Unity 3D

To import the skeleton tracking data from Kinect to Unity 3D, an updated AM Plugin was generated based on the previous version used in previous work [5] [31]. A dynamic link library (DLL) file was created and called by an AM Plugin Script in Unity 3D. Natural User Interface (NUI) was as a reference provided by Microsoft, which listed APIs for identifying the gesture as user input.

An AM plugin DLL file was generated to be called in Unity 3D via AM plugin script, which made Unity 3D interfacing with Kinect motion tracking sensor. An index of skeleton joints was introduced in the AM plugin script (Table 4.5).

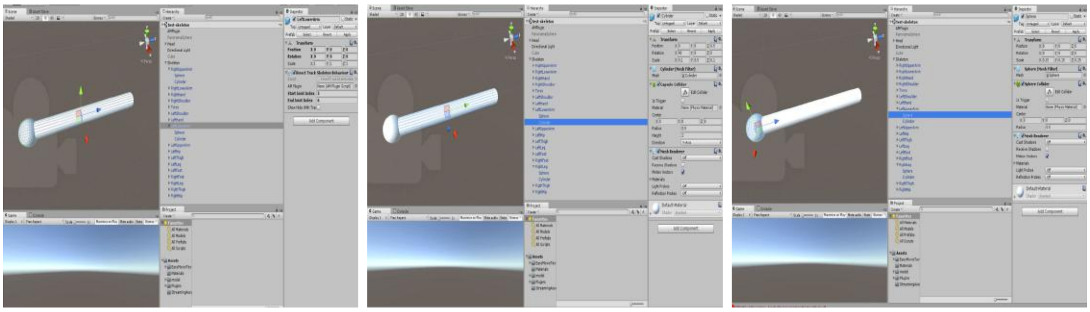
Table 4.5: Index of each joint defined in AM plugin script

Joint	**Index**
HIP CENTER	0
SPINE	1
SHOULDER CENTER	2
HEAD	3
SHOULDER LEFT	4
ELBOW LEFT	5
WRIST LEFT	6
HAND LEFT	7
SHOULDER RIGHT	8
ELBOW RIGHT	9
WRIST RIGHT	10
HAND RIGHT	11
HIP LEFT	12
KNEE LEFT	13
ANKLE LEFT	14
FOOT LEFT	15
HIP RIGHT	16
KNEE RIGHT	17
ANKLE RIGHT	18
FOOT RIGHT	19
SKELETON COUNT	20

Each joint has its index, which helped to configure the skeleton based cylinders and spheres in the Unity 3D.

The skeleton-based cylinders and spheres were created for every two joints in the Unity 3D environment. Therefore, these could be visualized by the user. Figure 4.14 shows the left lower arm skeleton with cylinder and sphere in Unity 3D. In the Kinect Track Skeleton Behaviour panel, the indexes were set from 5 to 6, which meant from elbow left joint to wrist left joint, indicating that the skeleton would follow the movement of the user's left lower arm.

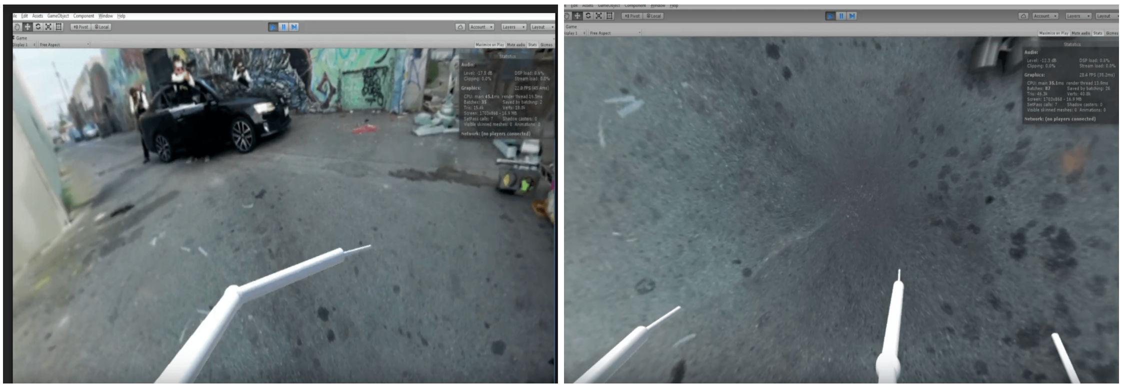
The skeletons are shown in Figure 4.15. The skeletons with cylinder and sphere were



(A) Left low arm skeleton (B) Left low arm cylinder (C) Left low arm sphere

FIGURE 4.14: Left low arm skeleton with cylinder and sphere in Unity 3D

generated to track the user's joints dynamically. This allows the virtual skeletons to be moved and showed in the 360° panoramic movie.



(A) (B)

FIGURE 4.15: User's skeleton shown in the 360° movie

4.4.3 3D Models augmented in Unity 3D

The previous section discussed the open source 3D Models in SketchUp. Figure 4.16 shows the completed 3D model of the armor.

This 3D armor consisted of significant components that did not match the skeletons set in Unity 3D. Therefore, 3D models were broken down before implementation in Unity 3D. Fortunately, the SketchUp Make provides this feature; the breakdown components of legs are shown in Figure 4.17. For instance, the initial 3D model only contained one component of the two legs (4.17A). However, there were four joints for each leg, including hip, knee, ankle, and foot. The researcher used an internal feature of the SketchUp to break down one component into six components, as shown in 4.17B.

The cylinders and spheres of the skeletons were already generated in Unity 3D. The components of the 3D model were implemented as being attached to the existing skeletons, which inherited from skeletons when running in the scene in Unity 3D. Figure 4.18

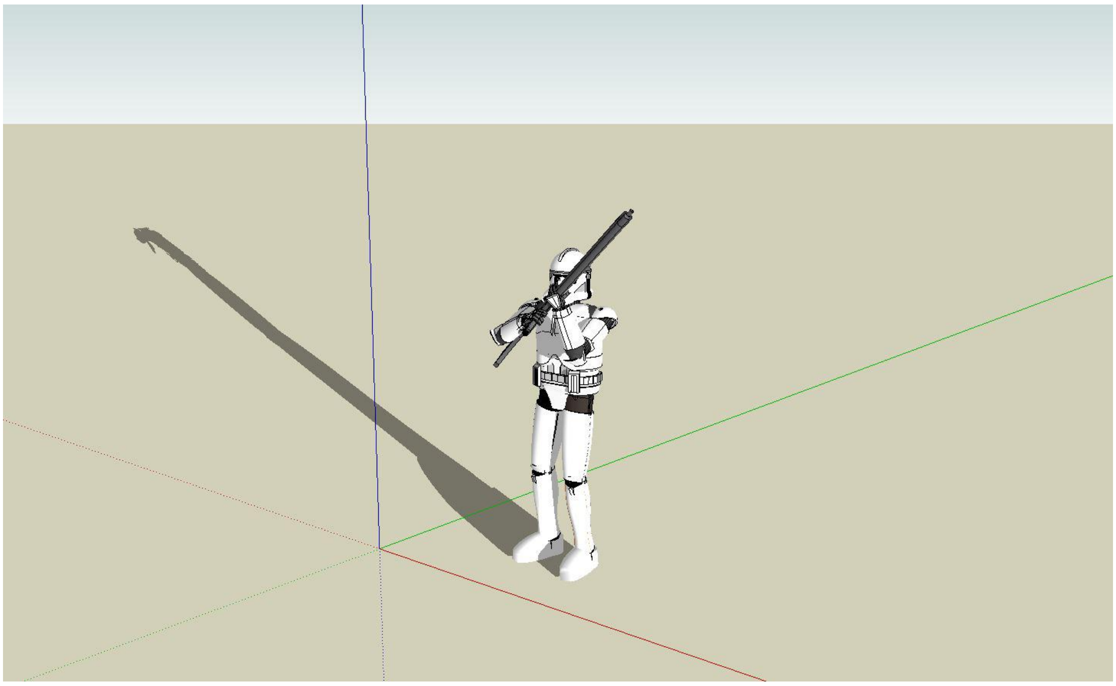
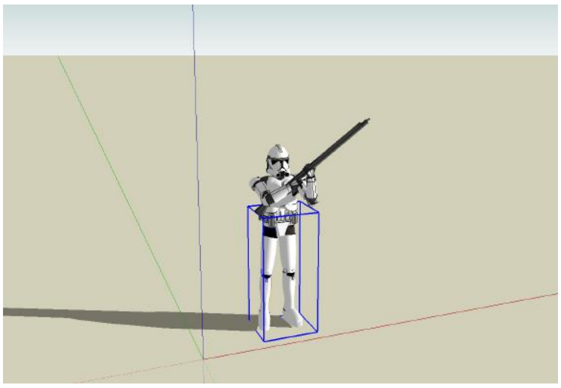
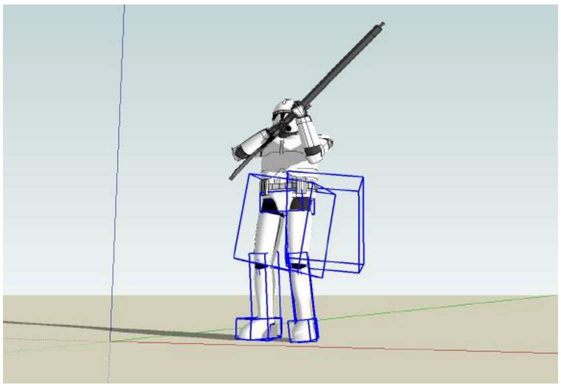


FIGURE 4.16: 3D armor set in SketchUp



(A) The initial leg's component of this armor



(B) The leg's components after breaking down

FIGURE 4.17: The components of armor's legs

presents the one piece of left leg component attached to the cylinder and sphere skeleton which combined left knee joint and left ankle joint.

The components were adjusted to match the required sizes and positions were aligned with existing skeletons. The features in Unity 3D provides a good solution to achieve calibrating, as shown in the process of calibrating the skeleton and armor's component in Figure 4.19.

3D models could be modified according to different themes and design requirements. The 3D models moved together with the skeleton based on user's skeleton tracking (Figure 4.20).

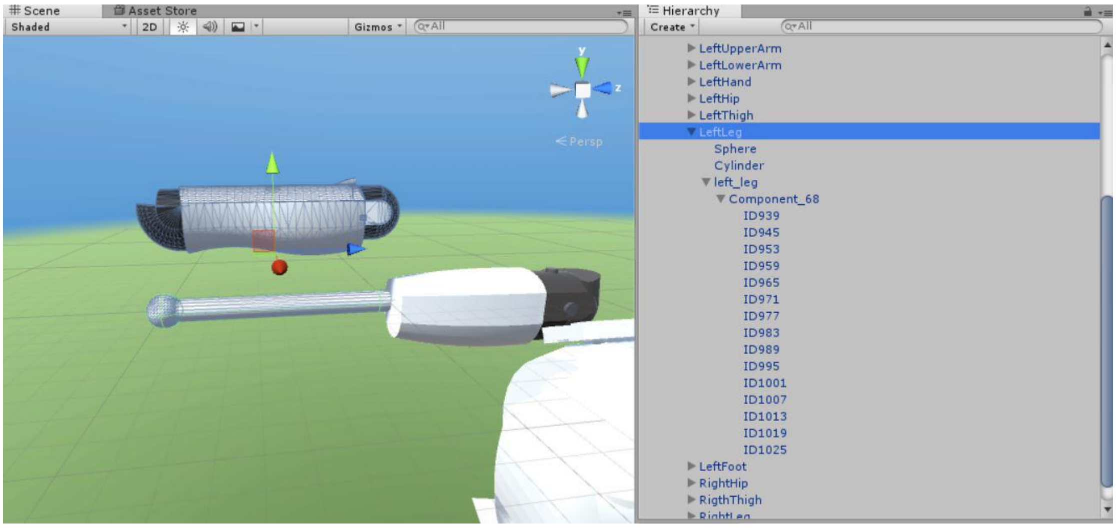
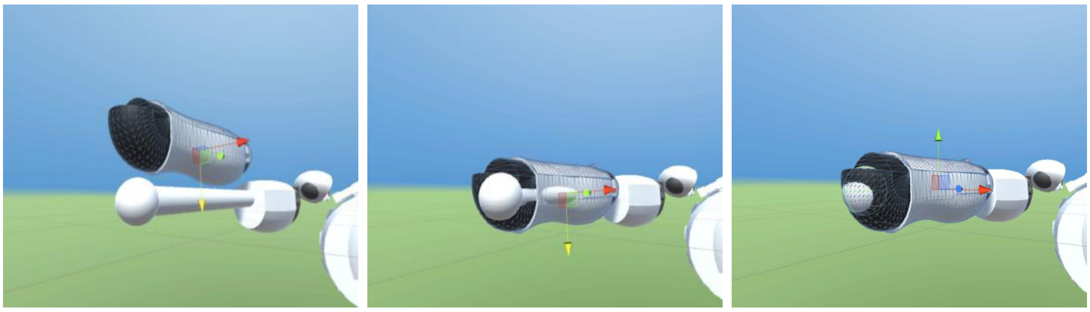


FIGURE 4.18: Left leg armor attached to left leg skeleton



(A) (B) (C)

FIGURE 4.19: Calibration of 3D armor's leg component and skeleton in Unity 3D

4.5 Content

4.5.1 3D Models

Limited to the capability of creating sound 3D armors and weapons, existing and free open source 3D models were chosen to be used in this project. The researcher initially thought of light clothing, armors, weapons, and tattoos which would be easily linked to the existing video content. After a comparison, as shown in Table 4.6, the researcher chose to use the armor and weapon in this project.

Different sets of armor and weapons can be obtained in SketchUp a 3D warehouse. Those models should be related to 360° movies. The armors are shown in Figure 4.21, while the weapons are shown in Figure 4.22.

Certainly, the Kinect skeleton tracking system has its limitations. After analyzing the effects of individual virtual weapons in Unity 3D, the gun had the better tracking result.

FIGURE 4.20: Left leg armor in running scene

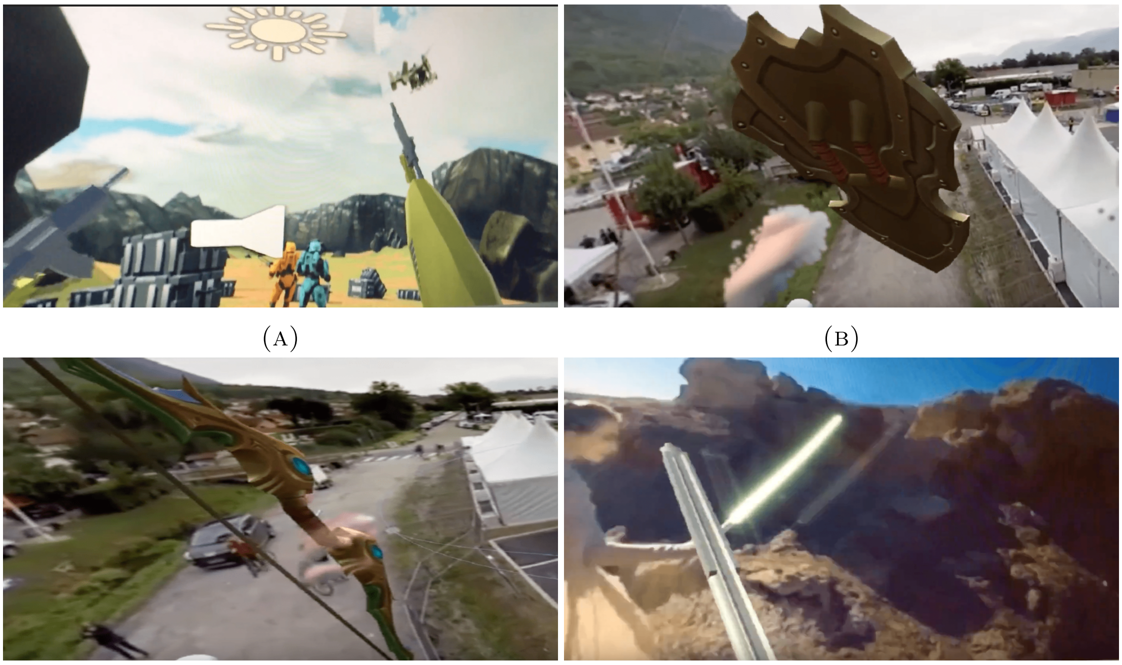
TABLE 4.6: The comparison of different categories of 3D model

Categories	**Advantages**	**Disadvantages**
Light Clothes	Colourful and attractive	Difficult to simulate the fluttering
Armor	Big enough to cover the skeleton	Looks like VR feature
Weapon	Easy to attract attention	User may expect interactions with it
Tattoo	Small, easy to stick to the body	Too small to be noticed

(A) (B)

FIGURE 4.21: Different models of armor

At this stage, the researcher only has limited choices for 3D models and 360° videos. The choice of 360° videos were set to match the existing 3D models, so that the weapon and armor would be well-contained in the scene. The type of armor had led to the choice of 360° videos being picked up from an open source.



(A) (B) (C) (D)

Figure 4.22: Different models of weapon

4.5.2 360° movie for immersive experience

Following from the discussion of the 3D models used in the project, it was concluded that the 360° movie used in this project should relate to armors and guns, which narrowed the search scope.

Initially, the researcher considered both realistic and animated background videos. Because armors and guns are widely used in animated games, a video “Red vs. Blue 360°: Supply Drop” [2] was chosen. The user would quickly adapt to the surroundings when they found themselves in the gun game environment; more interactions are expected to develop when all the surrounding characters are also wearing armor and are equipped with guns, as shown in Figure 4.23.

Regarding the realistic video, it would be more challenging to find armor and gun augmented content. Therefore, the study used on a timeless classic Star Wars themed video, which was a video “Star Wars 360° VR Experience — Desert Assault” [3], as shown in Figure 4.24.

In further consideration during the working process, the researcher decided to conduct the study with the realistic video only, which fitted the AV project better. It reduced

[2] https://www.youtube.com/watch?v=1wLCbAdLZUs
[3] https://www.youtube.com/watch?v=Ha40JjnU0wA&t=2s

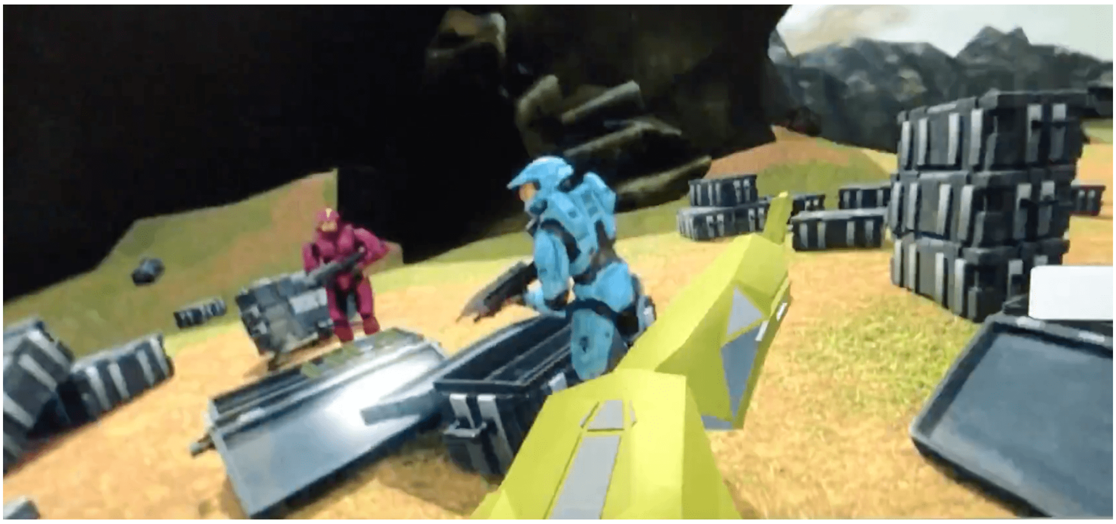


FIGURE 4.23: Red vs. Blue 360°: Supply Drop

FIGURE 4.24: Star Wars 360° VR Experience — Desert Assault

the AV effects in the animated video, which the users sometimes confused with the VR effects.

4.5.3 Enhanced Visualization Prototype

An enhanced visualization working prototype was generated for user experience. The video was trimmed to an appropriate two-minute length. To fit the Star Wars theme, the characters in the scene were armed with Stormtrooper Armor in a desert.

Three different scenes had been created in Unity 3D, representing three different visualizations. The first was the " movie only" scene, which played the 360° movies only in the Unity, as shown in Figure 4.25A. The second was the "movie with real body" scene, which blended the user's real body in the scene. If the user moved their hands or legs within the filed of view of the SoftKinetic camera, they could see their real body in the

(A) The first scene (B) The second scene (C) The third scene

FIGURE 4.25: User's visualization in three different scenes

360° movie scene, as shown in Figure 4.25B. The third was the "movie with the real body overlaid by virtual objects". The real body inside the SoftKinetic camera field of view would be shown in the scene, while the Kinect tracked the user's movements that were overlaid by a virtual white Stormtrooper armor and a gun, as shown in Figure 4.25C. The virtual armors were well matched to the character's armors, as well as the gun.

The enhanced visualization in the third scene was the principal scene this study would like to test. In the 360° video, the user could move their body in any way they would like, and the Kinect would track their skeleton to import data into Unity. In the Oculus Rift DK2 headset, the user could see their real body overlaid with by the Stormtrooper armor in the field of view of the Softkinetic camera.

The screenshots captured from Unity "movie with the real body overlaid by virtual objects" scene are shown in Figure 4.26. The user could watch the 360° movie from any angle using Oculus Rift DK2 headset, and visualize their real body in a similar Stormtrooper armor.

4.5.4 Trial Run

A trial run was conducted with two participants to acquire feedback after finalizing the working prototype.

The armor and gun were augmented on the right or left hands skeleton, while the real hand was rendered by the SoftKinetic camera. The user could move their hands or legs gently, and the theme based armor and weapon would follow their movements.

When the researcher ran the scene, one user had a significant alignment issue. The alignment was a certain distance away from the correct position. The researcher investigated

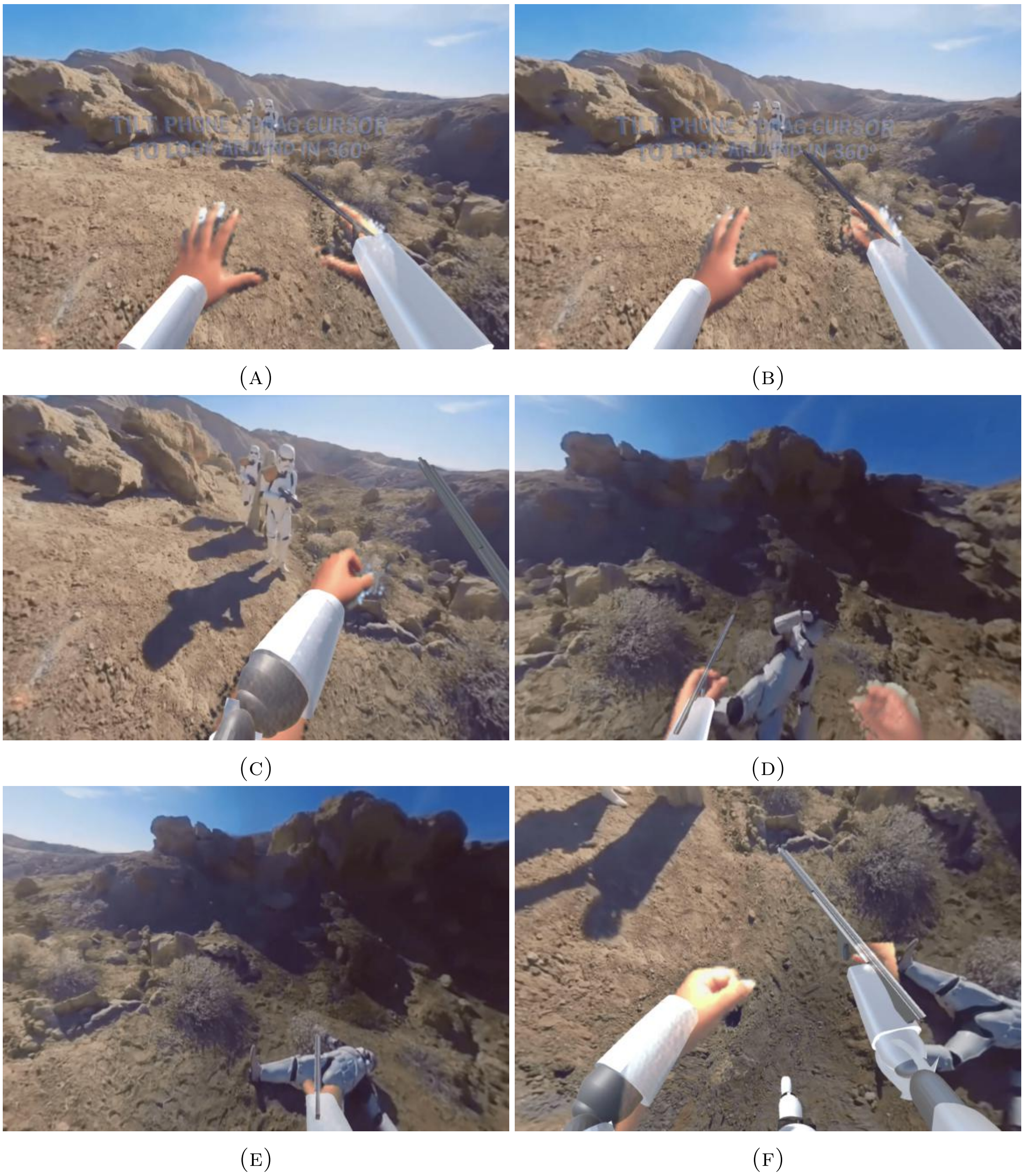


FIGURE 4.26: User's visualizations in the "movie with the real body was overlaid by virtual objects" scene

the issue and found that the problems lies with the alignment between Oculus Rift head tracker sensor and the Kinect head joint. The user had occasionally looked around the environment while the application initialized, which set the displacement between the Oculus Rift head tracker sensor and the Kinect, as shown in Figure 4.27. Therefore, one extra requirement was added for users: look straight at the Kinect and stay still for a few seconds when the application is initializing.

(A) Displacement before head alignment

(B) Fixed after head alignment

FIGURE 4.27: Before and after head alignment between Kinect and Oculus Rift head tracker

4.6 Summary

This chapter explained the implementation of the working prototype. Firstly, the hardware setup was described including the Kinect, the Oculus Rift DK2 headset and sensor, the SoftKinetic DepthSense 325 camera, the headphone and the computer. Secondly, the software part had been updated, the key AM plugin script was called in the Unity to import data, and the movie texture plugin script was used to generate objects to play the 360° movie. Moreover, virtual objects were augmented on the existing skeleton tracking to simulate the movie's character and content which had been carefully chosen. A working prototype had been generated in the Unity 5 game engine, including three scenes in the project. Finally, after the working prototype had been finalized, a trial run was conducted to test the prototype. The next chapter will describe the user experiment which is conducted based on the working prototype described in this chapter.

Chapter 5

User Evaluation

This chapter describes the user evaluation of the working prototype, including experiment design, hypotheses, experiment procedures and tasks.

5.1 Evaluation Goal

An evaluation is concerned with gathering data about the quality of a design or product. There several reasons the researcher conducts the evaluation [48].

- To validate the prototype
- To refine our prototype and design
- To learn more about user and the problem
- To move forward to the next iteration

The methods of evaluation in Human-Computer Interaction (HCI) research are observation, ethnography, case study, interview, focus group, survey, questionnaire, usability testing and user experiment. The methods used range from informal, qualitative feedback and formative testing to formal, quantitative feedback, summative and validation testing [48]. The outcome of the evaluation is the user feedback, which would affect the design idea for the prototype (Figure 5.1).

Providing a compelling, immersive cinematic experience is the critical goal for producers of 360° movies. The evaluation is an essential section of the research for user experience. This chapter will describe the experimental design to achieve the goal of the experiment. The information sheet and consent form used in the experiment as well as the approval

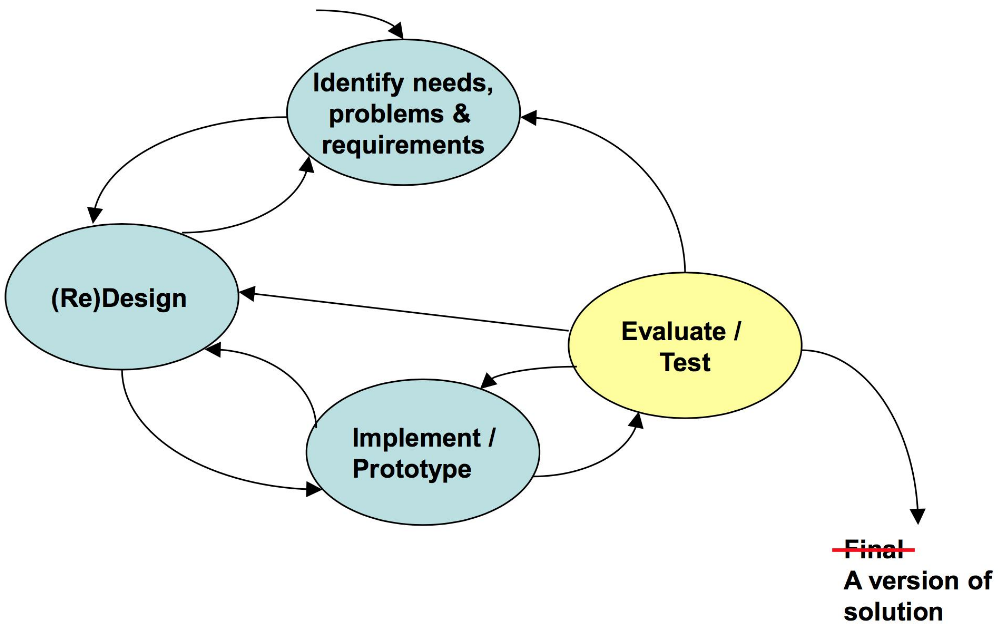


FIGURE 5.1: Human-Computer Interaction research and evaluation [48]

of the user experiment from the human ethics office at the University of Canterbury are found in Appendix A. Appendix B lists all the related subjective questionnaires the researcher used for the user experiment, including a pre-experiment questionnaire, per-condition questionnaire for each condition, and a post-experiment questionnaire.

The hypothesis for the final experiment is that there is a significant improvement of user experience by augmenting the blended physical objects in the movie scene with the theme matching the virtual content. A survey with Likert-scale rating questions runs as a measurement in the user experiment, to assess whether augmenting AV improves the sense of presence for user experience.

5.2 User Experiment Design

User experiment is a method of academic research in HCI, trying to test new technology. It is hypothesis-driven, comparing multiple conditions, which can be replicated. After the experiment, results are drawn using statistical analysis of the obtained data.

5.2.1 Variables and Conditions

The independent variable for the user experiment has three levels, namely “Movie Only”, “Real Body”, and “context-aware objects and real body”. The factorial design for the experiment is shown in Table 5.1.

Factor	**Movie Only**	**Real body**	**Context-aware objects and real body**
360° movie	A	B	C

TABLE 5.1: Experiment design

All possible permutations are used in the experiment to counter-balance the order effect. Three conditions which means six permutations in total, see Table 5.2.

A	B	C
A	C	B
B	A	C
B	C	A
C	A	B
C	B	A

TABLE 5.2: All possible permutations design

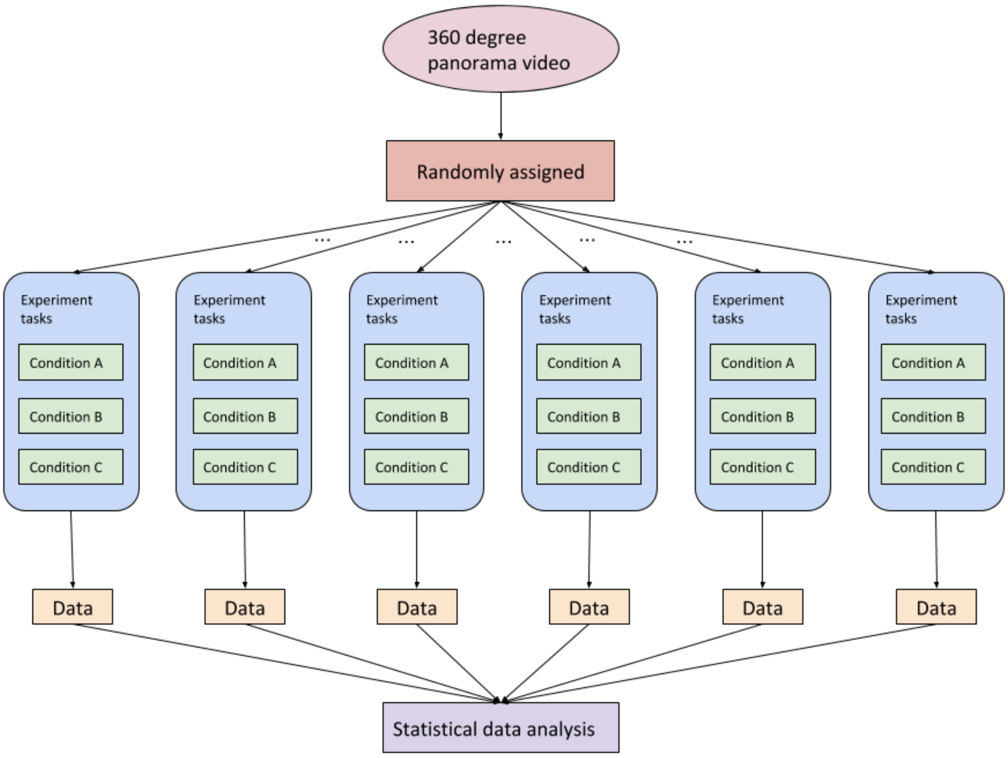


FIGURE 5.2: Within-subject experimental design

The experiment consists of three conditions. Condition A is the control, which only includes visualizing a movie, condition B is visualizing user's real body in the movie, and condition C is visualizing user's real body with context-aware virtual objects in the 360° movie. The user experiment uses a within-subjects experimental design, where participants would attempt all three conditions in random orders, as shown in Figure 5.3.

The experiment used the Igroup Presence Questionnaire (IPQ) [1] which is designed and created by Igroup to test for the sense of presence in a virtual environment.

5.2.2 Hypotheses

We have two hypotheses in this experiment.

- Hypothesis H_1: There is a significant difference in the sense of presence of enhanced visualization between visualizing a movie only, visualizing user's real body and visualizing context-aware virtual objects augmented on user's real body in a cinematic scene when experiencing the 360° movie using HMDs.

- Hypothesis H_2: There is a significant difference in the user preference of enhanced visualization between visualizing a movie only, visualizing user's real body and visualizing context-aware virtual objects augmented on user's real body in a cinematic scene when experiencing the 360° movie using HMDs.

5.2.3 Experiment Setup

The experiment was conducted at the Student Lab in the Human Interface Technology (HIT) Lab NZ, University of Canterbury. The setup is shown in Figure 5.3.

The items the researcher used for this experiment were the Kinect, an Oculus DK2, a SoftKinetic DepthSense 325 camera, a headphone and a laptop/desktop, as detailed in Chapter 4. Each participant was assigned all three conditions in random order. The 360° movie "Star Wars 360° VR Experience — Desert Assault" had been trimmed to about 2 minutes, to improve participants engagement in the experiment.

5.2.4 Experiment procedure

The whole experiment took approximately 35-40 minutes and followed the below procedures.

[1] http://www.igroup.org/pq/ipq/

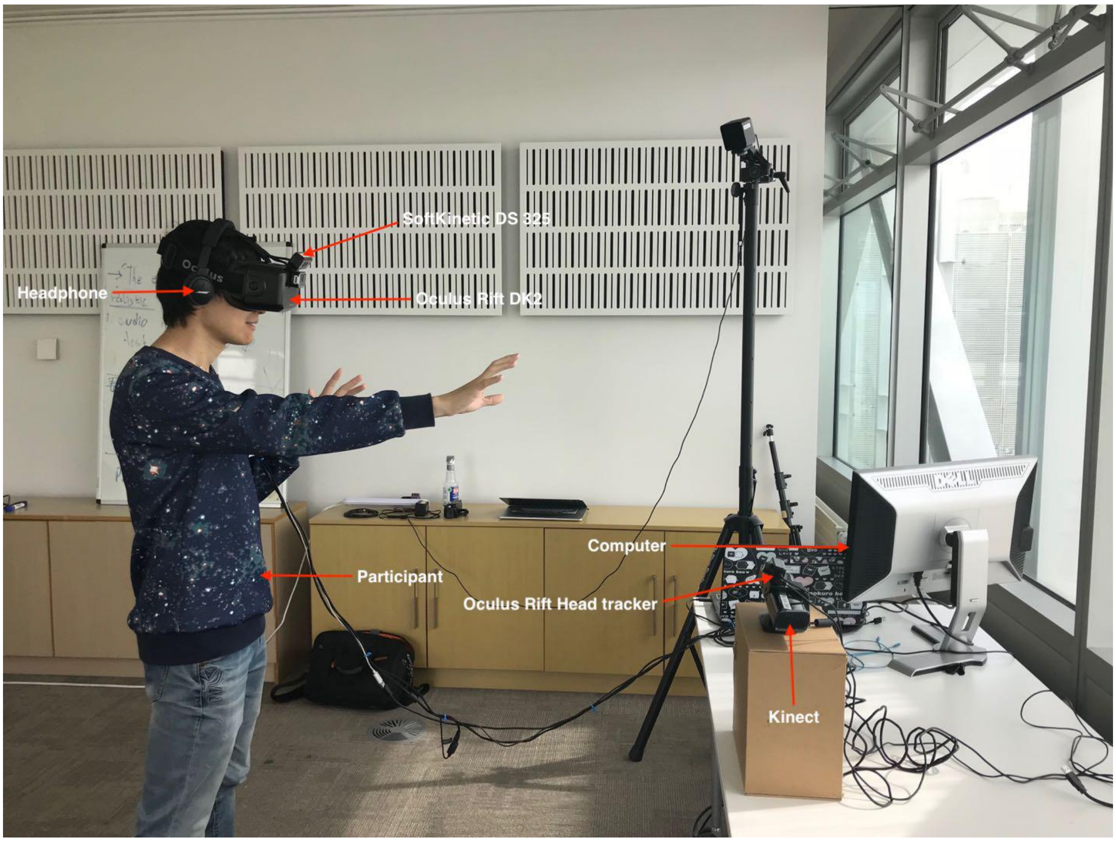


FIGURE 5.3: Experiment Setup

1. The participants were given a full explanation about the general structure of the project at the beginning of the experiment, and they were given the opportunity ask any questions regarding the project. Then, the information sheet and the consent form (Appendix A) were provided for them to read and sign.

2. The participants were then given a pre-experiment questionnaire (Appendix B) to gather demographic information and their previous experience with using HMD.

3. A brief explanation of the experimental setup and tasks were provided to the participants, as well as the experimental process.

4. The participants were given the working prototype, an Oculus Rift DK2 HMD with SoftKinetic camera mounted on, and a headphone to wear. The researcher initialized the devices once the participants were ready and aligned the head tracker with Kinect sensor at the start of each condition.

5. The participants would then experience all three conditions in a sequence defined by all possible permutations to perform experimental tasks. A per-condition questionnaire (Appendix B), is given to fill at the end of each condition.

6. Once the participant finished all three conditions, a post-experiment questionnaire (Appendix B) was dispatched to the participant regarding the overall experience and feedback.

7. Finally, participants were given a short debriefing session to identify any questions or suggestions regarding the experimental process and tasks.

5.2.5 Experiment Task

The participants were required to experience all three conditions, and their tasks were list below (Figure 5.4).

Condition A: The participant watches the 2 minutes 360° movie "Star Wars 360° VR Experience — Desert Assault" without any enhanced visualization in the movie.

Condition B: While watching the 2 minutes 360° movie "Star Wars 360° VR Experience — Desert Assault", the participants were encouraged to move their hands and body to see their real body blended in the movie scene.

Condition C: While watching the 2 minutes 360° movie "Star Wars 360° VR Experience — Desert Assault", the participants were encouraged to move their hands and body to see context-aware virtual objects augmented on the real body in the movie scene.

(A) Condition A (B) Condition B (C) Condition C

FIGURE 5.4: Experiment task for each condition

5.2.6 Measures

Sense of presence was the principal measure for the experiment. The Igroup Presence Questionnaire [2] was used to measure the user's sense of presence experienced in a VE. It contained four main components, general presence (PRES), spatial presence (SP), involvement (INV) and experienced realism (REAL).

[2] http://www.igroup.org/pq/ipq/download.php

The scale development process has identified one general item for general presence, five items for spatial presence, four items for involvement, and three items for experienced realism, as shown below [49].

- General Presence (PRES) : describe "sense of being there" [49];
- Spatial presence (SP): subjective measurement of the own experiences of presence [49];
- Involvement (INV): measure the attention to the real and virtual environment [49];
- Experienced Realism (REAL): measure the subjective experience of a comparison between the virtual and the real world [49].

The IPQ was used in our experiment to measure the sense of presence for all three conditions.

5.2.7 Pilot Study

A pilot study was conducted before the formal experiment, which invited two participants, one female and one male. They went through the entire experiment, including all conditions and questionnaires, similar to the formal experiment. The pilot study was conducted to identify any issues on the procedures and the prototype. The data were collected and recorded, but not included in the analysis.

In our pilot study, the researcher resolved a configuration issue which caused system crashes, due to a compatibility setting between 32-bit and 64-bit for AM plugin DLL file. The system crash issues was fixed and not shown again in the formal experiment.

5.3 Summary

This chapter mainly explained the user experiment design, tasks, procedures and hypotheses. The user experiment was conducted to evaluate the working prototype described in Chapter 4. The results of this user experiment will be discussed in the next chapter.

Chapter 6

Results

This chapter reports the user experiment results from an analysis on the data collected from 24 participants. Each participant completed:

- one pre-experiment questionnaire for demographic information
- three per-condition questionnaires for quantitative data for each condition
- one post-experiment questionnaire for quantitative and qualitative data to evaluate user's overall experience.

6.1 Demographics

Twenty-four participants were recruited after obtaining the approval letter from the University human ethics committee (Appendix A).

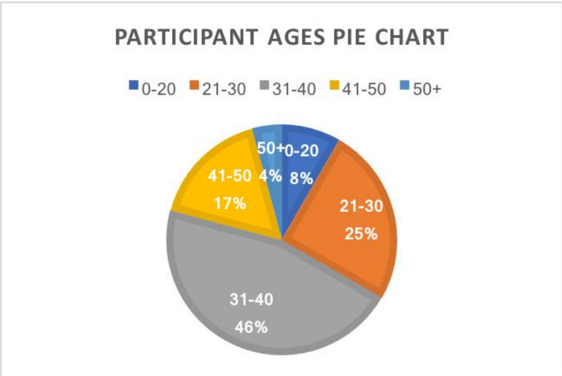


FIGURE 6.1: Participant ages pie chart

There were 12 (50%) female participants and 12 (50%) male participants. The ages ranged from 13 to 52 (Figure 6.1), with an average of 32.42, and a standard deviation of 8.59. Nearly half (46%) of the participants were in their 30s.

The bar chart in Figure 6.2 illustrates participants' frequencies of using HMD and watching 360° movies. As shown in the diagram, eight (33.3%) participants had never used a HMD and 11 (45.8%) had never watched 360° movies before. Two (8.3%) participants were very familiar with the HMD with daily or weekly frequency of use.

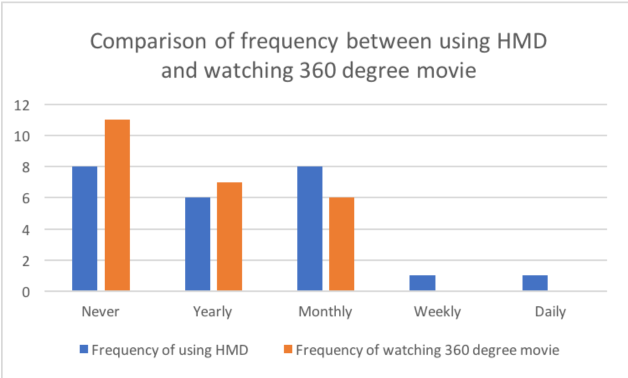


FIGURE 6.2: Comparison of participants' frequency between using HMD and watching 360° movies

A clustered bar chart had been generated to show the initial opinion of seeing real body and virtual costumes in the movie (Figure 6.3). This diagram offered an overview of how people think of the initial ideas of the experiment before experiment. These two were set to seven-point Likert scale question. There were nine (37.5%), and five (20.8%) participants had a neutral option (rating = 0) of seeing the real body and seeing the real body with virtual costumes. Most of the participants had a favourable opinion (rating >0) about seeing the real body and seeing the real body with virtual costumes in the movies. Only two (8.3%) and one (4.2%) participants had a negative option (rating <0) about seeing the real body with or without the virtual costumes.

6.2 Quantitative Measures

There is one independent variable with three levels in the experiment. The dependent variable is the participants' sense of presence, which is collected from the questionnaires after three conditions. The per-condition questionnaire consists of a set of questions with a seven-point Likert scale, which was treated as ordinal data while the non-parametric test was applied in this experimental study. A non-parametric test was conducted to compare the ordinal measure. Friedman test was used as three conditions against expected value using within-subject design. This test was followed by the post-hoc tests,

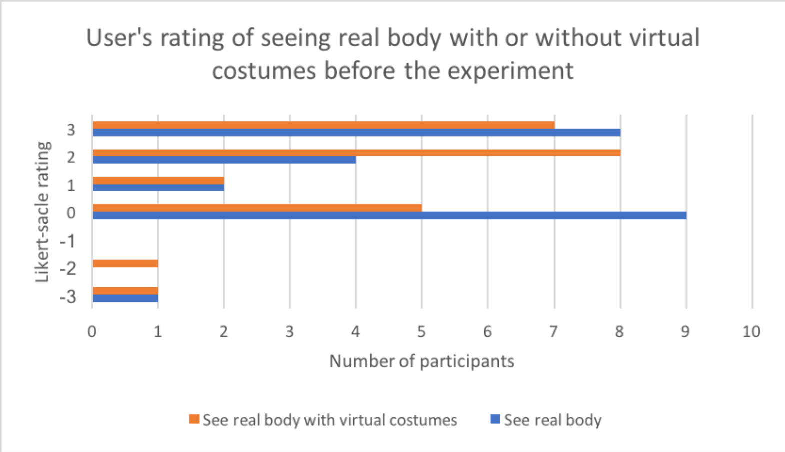


FIGURE 6.3: The initial opinion of seeing real body and avatar in the movie

which was a set of Wilcoxon signed rank tests using Bonferroni correction making the pair-wise comparison. The significant level of Friedman test would be $\alpha = 0.05$. Because of three three pair-wise comparison, the αvalue in Wilcoxon signed rank tests was adjusted with Bonferroni correction, $\alpha = 0.05 / 3 = 0.017$.

Moreover, the post-experiment rank question was considered to be an ordinal data and analysed with Friedman test, and Wilcoxon signed rank tests.

6.2.1 Igroup Presence Questionnaire,IPQ

IPQ has 14 seven-point Likert-scale items, which are rated from zero to six. The researcher analyzed the IPQ, as well as the four components of IPQ. The first 14 questions of per-condition questionnaire are the items of IPQ questionnaire. Question 8 is classified as PRES item. Question 9, question 13, question 6, question 3,and question 10 are classified as SP. Question 1, question 7, question 11, and question 14 are classified as INV. Question 2, question 4, question 5 and question 12 are classified as REAL. According to the guide from IPQ, three items had been reversed with the rating value before they were calculated. These three items are Question 13, 11 and 2. There are three conditions for analyzing, condition A is visualizing a movie only, condition B is visualizing user's real body in the movie, condition C is visualizing user's real body with context-aware virtual objects in the movie.

6.2.1.1 General Presence, PRES

The PRES is a component of the IPQ questionnaire. It consists of one question. We applied Friedman test followed by post-hoc tests using the Bonferroni correction and generate the box plot. There was a statistically significant difference between condition

A, condition B and condition C, $\chi^2(2) = 10.099$, $p = 0.006$. Median PRES levels for condition A was 3.00, condition B was 4.00 condition C was 4.00, respectively see in Table 6.1 and 6.2.

TABLE 6.1: General presence descriptive statistics table

GP

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Descriptive Statistics

		Percentiles		
	N	25th	50th (Median)	75th
Condition A	24	1.00	3.00	4.00
Condition B	24	2.25	4.00	4.75
Condition C	24	3.00	4.00	5.00

TABLE 6.2: General presence Friedman test result

Friedman Test

Ranks

	Mean Rank
Condition A	1.52
Condition B	2.17
Condition C	2.31

Test Statistics[a]

N	24
Chi-Square	10.099
df	2
Asymp. Sig.	.006

a. Friedman Test

Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. There were a statistically significant improvement in sense of PRES between condition A and condition B ($Z = -2.743$, $p = 0.006$) or between condition A and condition C ($Z = -2.850$, $p = 0.004$).

TABLE 6.3: General presence Wilcoxon signed-rank result

Wilcoxon Signed Ranks Test

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Test Statistics[a]

	Condition B – Condition A	Condition C – Condition A	Condition C – Condition B
Z	-2.743[b]	-2.850[b]	-.292[b]
Asymp. Sig. (2-tailed)	.006	.004	.770

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

However, there were no significant differences between condition B and condition C (Z = -0.292, p = 0.770), as shown in Table 6.3.

A box-plot was generated for the general presence scores, as shown in Figure 6.4.

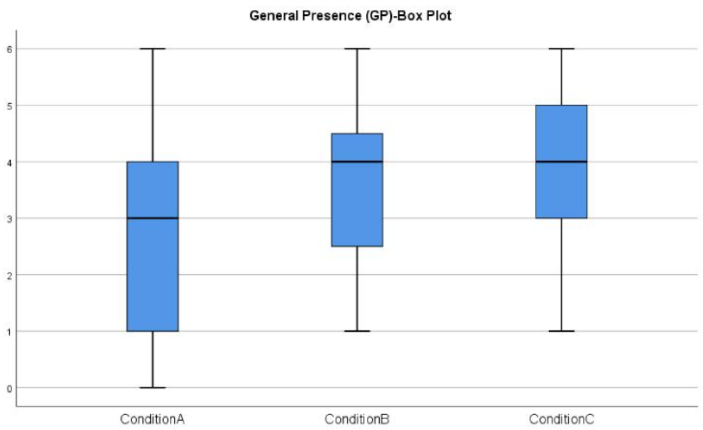


FIGURE 6.4: General presence Box-plot

6.2.1.2 Spatial Presence, SP

Figure 6.5 describes the box-plot for spatial presence.

In SP, the researcher applied Friedman test followed by post-hoc tests using the Bonferroni correction and generate the box plot. There was a statistically significant difference between condition A, condition B and condition C, $\chi^2(2) = 9.023$, p = 0.011. Median

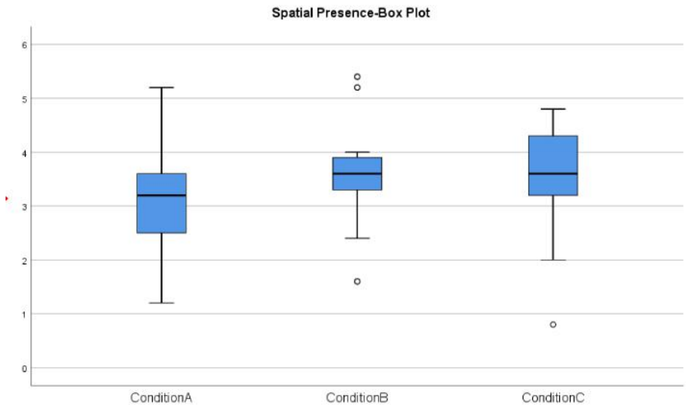


Figure 6.5: Spatial presence Box-plot

Table 6.4: Spatial presence descriptive statistics table

SP

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Descriptive Statistics

	N	Percentiles		
		25th	50th (Median)	75th
Condition A	24	2.450	3.200	3.550
Condition B	24	3.250	3.600	3.950
Condition C	24	3.400	3.700	4.600

SP levels for condition A only was 3.20, condition B was 3.60 while condition C was 3.70 respectively, as shown in Table 6.4 and 6.5.

Post hoc analysis with Wilcoxon signed-rank tests was conducted with a Bonferroni correction applied, resulting in a significance level set at $p < 0.017$. There was a statistically significant improvement in the sense of general presence between condition A and condition C ($Z = -2.993$, $p = 0.003$). However, there were no significant differences between condition A and condition B ($Z = -2.315$, $p = 0.021$) or between condition B and condition C ($Z = -0.714$, $p = 0.475$), as shown in Table 6.6.

TABLE 6.5: Spatial presence Friedman test result

Friedman Test

Ranks

	Mean Rank
Condition A	1.54
Condition B	2.13
Condition C	2.33

Test Statistics[a]

N	24
Chi-Square	9.023
df	2
Asymp. Sig.	.011

a. Friedman Test

TABLE 6.6: Spatial presence Wilcoxon signed-rank result

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Test Statistics[a]

	Condition B – Condition A	Condition C – Condition A	Condition C – Condition B
Z	-2.315[b]	-2.993[b]	-.714[b]
Asymp. Sig. (2-tailed)	.021	.003	.475

a. Wilcoxon Signed Ranks Test

b. Based on negative ranks.

6.2.1.3 Involvement, INV

In INV, the researcher applied Friedman test. There was no statistically significant difference between condition A, condition B and condition C, $\chi^2(2) = 0.945$, $p = 0.623$. Median INV levels for condition A was 3.250, condition B was 3.375 while condition C was 3.250 respectively, as shown in Table 6.7 and Table 6.8.

TABLE 6.7: Involvement descriptive statistics table

INV

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Descriptive Statistics

	N	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Condition A	24	2.2500	3.2500	4.0000
Condition B	24	2.0000	3.3750	4.0000
Condition C	24	2.5625	3.2500	3.5000

TABLE 6.8: Involvement Friedman test result

Friedman Test

Ranks

	Mean Rank
Condition A	2.13
Condition B	1.85
Condition C	2.02

Test Statistics[a]

N	24
Chi-Square	.945
df	2
Asymp. Sig.	.623

a. Friedman Test

The box-plot of involvement is shown in Figure 6.6.

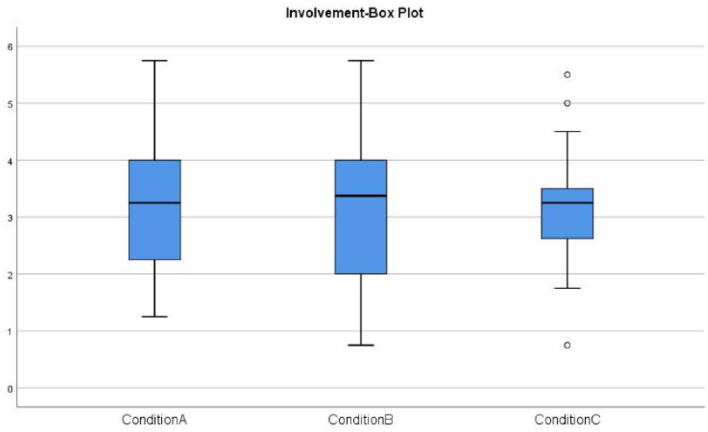


Figure 6.6: Involvement Box-plot

6.2.1.4 Experienced Realism, REAL

The same method had been implemented in the REAL component. Based on the Friedman test, there was no statistically significant difference between condition A, condition B and condition C, $\chi^2(2) = 1.744$, $p = 0.418$. Median REAL levels for condition A was 2.63, condition B was 2.75 while condition C was 2.88 respectively, as shown in Table 6.9 and Table 6.10.

Table 6.9: Experienced realism descriptive statistics table

REAL

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Descriptive Statistics

		Percentiles		
	N	25th	50th (Median)	75th
Condition A	24	2.2500	2.6250	3.1875
Condition B	24	2.5000	2.7500	3.1875
Condition C	24	2.3125	2.8750	3.5000

The diagram in Figure 6.7 states the box-plot of three conditions in experienced realism.

Table 6.10: Experienced realism Friedman test result

Friedman Test

Ranks

	Mean Rank
Condition A	1.79
Condition B	2.10
Condition C	2.10

Test Statistics[a]

N	24
Chi-Square	1.744
df	2
Asymp. Sig.	.418

a. Friedman Test

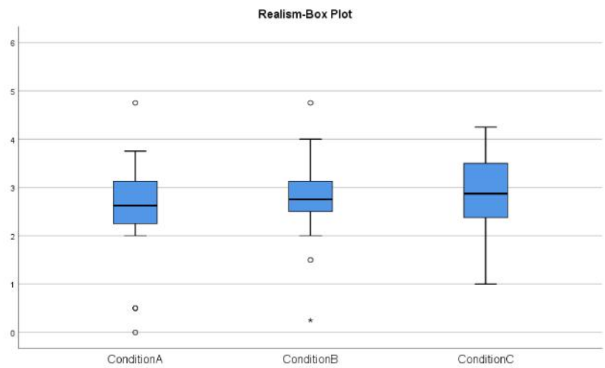


Figure 6.7: Experienced realism Box-plot

6.2.1.5 Overall IPQ

The overall IPQ is calculated by all 14 items. The previous section has already described the categorizes of these items. Question 8 was named G1. Question 9, question 13, question 6, question 3 and question 10 were named as SP1, SP2, SP3, SP4 and SP5. Question 1, question 7, question 11, and question 14 were named as INV1, INV2, INV3 and INV4. Question 2, question 4, question 5 and question 12 were named as REAL1, REAL2, REAL3 and REAL4.

The scale point is from zero to six, zero states the worse general/ presence/ involvement/ real, while six means better presence involvement/ real in most items. However, there are three items with reversed wording: SP2, INV3, and REAL1.

A Friedman test was conducted on the data collected from 24 participants. Table 6.11 shows the descriptive statistics, while Table 6.12 shows the rank table and the test statistics table.

TABLE 6.11: Overall IPQ descriptive statistics table

IPQ

Condition A = movie only
Condition B = movie with rendered real body
Condition C = movie with rendered context aware virtual object augmented on real body

Descriptive Statistics

		Percentiles		
	N	25th	50th (Median)	75th
Condition A	24	2.5536	3.0000	3.4286
Condition B	24	2.7321	3.3571	3.7500
Condition C	24	2.6607	3.4286	3.9821

There was no statistically significant difference between condition A, condition B and condition C, $\chi_2(2) = 3.095$, p = 0.213.

Median IPQ levels for condition A was 3.00, condition B was 3.36 while condition C was 3.43, respectively.

The overall IPQ scores box plot is shown in Figure 6.8.

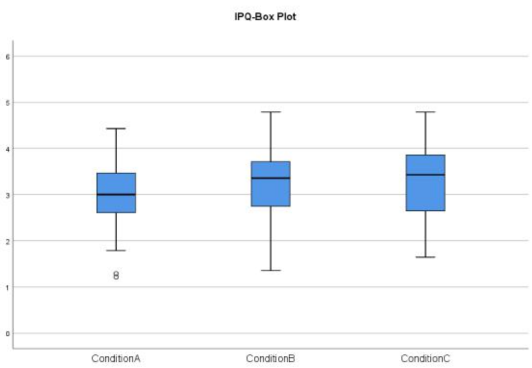


FIGURE 6.8: Overall IPQ Box-plot

TABLE 6.12: Overall IPQ Friedman test result

Friedman Test

Ranks

	Mean Rank
Condition A	1.73
Condition B	2.04
Condition C	2.23

Test Statistics[a]

N	24
Chi-Square	3.095
df	2
Asymp. Sig.	.213

a. Friedman Test

6.2.2 Post-experiment Questionnaire

Once the participants completed all conditions, a post-experiment questionnaire was given to collect feedback on their overall experience.

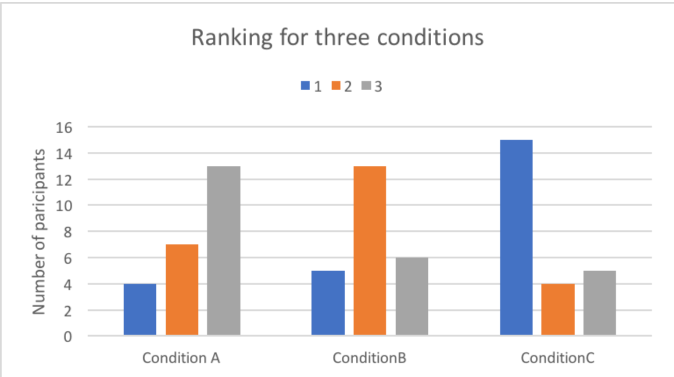


FIGURE 6.9: Ranking bar chart

A ranking question from one (best) to three (worst) was offered, the results are shown in the bar chart in Figure 6.9. Condition A represented watching a movie only, condition B represented watching a movie with the rendered real body, while condition C represented watching a movie with the rendered context-aware virtual object augmented on the real body.

TABLE 6.13: Friedman test result for ranking

Ranks

	Mean Rank
Condition A	2.38
Condition B	2.04
Condition C	1.58

Test Statistics[a]

N	24
Chi-Square	7.583
df	2
Asymp. Sig.	.023

a. Friedman Test

TABLE 6.14: Wilcoxon signed rank tests result SPSS

(A) Descriptives table in Wilcoxon signed rank test SPSS

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles 25th	Percentiles 50th (Median)	Percentiles 75th
Condition A	24	2.38	.770	1	3	2.00	3.00	3.00
Condition B	24	2.04	.690	1	3	2.00	2.00	2.75
Condition C	24	1.58	.830	1	3	1.00	1.00	2.00

(B) Test statistics table

Test Statistics[a]

	Condition B - Condition A	Condition C - Condition A	Condition C - Condition B
Z	-1.292[b]	-2.475[b]	-1.527[b]
Asymp. Sig. (2-tailed)	.197	.013	.127

a. Wilcoxon Signed Ranks Test

b. Based on positive ranks.

A Friedman test had been conducted as well, a non-parametric test for within-subject three conditions, as shown in Table 6.13. There was a statistically significant difference in three conditions, $X^2(2) = 7.583$, $p = 0.023$. We then performed a post hoc Wilcoxon signed rank tests with a Bonferroni correction applied, making the pair-wise comparison. We found a significance level set at $p < 0.017$. Mean rank for condition A was 2.38, condition B was 2.04 and condition C was 1.58. There was a statistically significant difference between condition A and condition C (Z = -2.475, p = 0.013). However, there were no significant differences between condition A and condition B (Z = -1.292, p = 0.197) or between condition B and condition C (Z = -1.527, p = 0.127), as shown in Table 6.14.

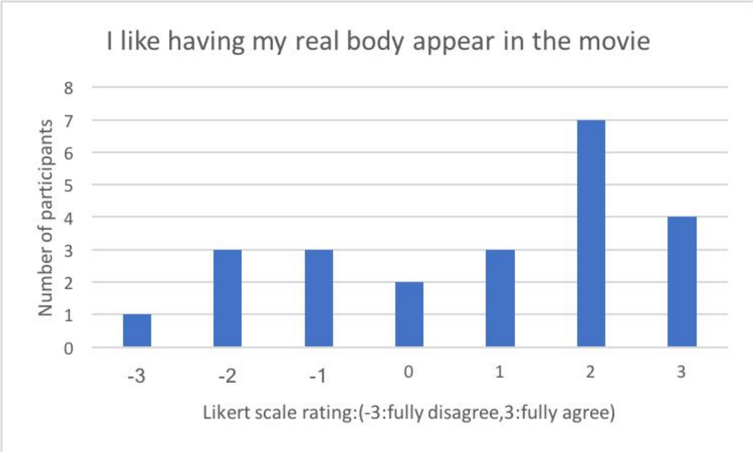


FIGURE 6.10: Participants rating of seeing their real body in a 360° movie

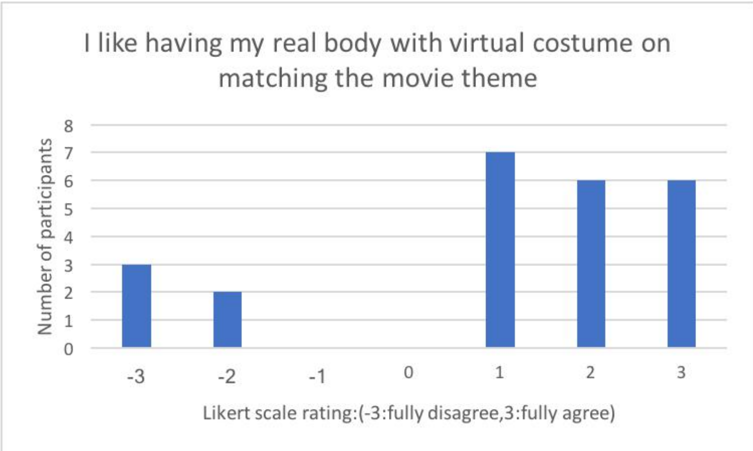


FIGURE 6.11: Participants rating of seeing their real body with virtual costume matching movie theme in a 360° movie

Participants were also asked whether they would like to see their real body or real body with theme content based virtual costume in the 360° movie, rating on a seven-point Likert-scale. Negative three represented fully disagree, while three meant fully agree. The results are shown in Figure 6.10 and Figure 6.11.

Table 6.15 describes the descriptive data of these two questions. The mean value were 0.83 and 1.00, while the median values were both 1.50.

TABLE 6.15: Descriptive Statistics table

	I like having my real body appear in the movie	I like having my real body with virtual costume on matching the movie theme
Mean	0.833333333	1
Min	-3	-3
Q1	-1	1
Median	1.5	1.5
Q3	2	2.25
Max	3	3

6.2.3 Quantitative Measures Summary

From the quantitative data, visualizing a movie only, visualizing user's real body in the movie and visualizing user's real body with context-aware virtual objects in a 360° movie had a significant effect on two presence components, general presence, and spatial presence.

Post-experiment questionnaire testing also described the significant difference in three conditions for ranking data, especially between condition A and condition C.

6.3 Qualitative Measures

6.3.1 Post-experiment Questionnaire

Apart from objective quantitative measures, questions in the post experiment questionnaire were also analyzed, to get further information regarding participants' preference for the experiment.

Participants were asked to explain their thoughts for ranking the three conditions.

Some participants had a positive feedback overall, either felt more involved or more interesting.

1. "Seeing a relatable virtual costume makes you feel more of being part of the movie."
2. "Real body with the virtual costume is more realistic."
3. "Personally prefer more involved in the movie rather than 3D picture only."

4. "Being in a costume made it feel more like I was a part of the movie rather than just watching it."
5. "Feel connected to the movie when having the virtual costume, but feel distracted only to see my body without the costume."
6. "New experience, interesting to be part of a movie."
7. "It feels like I am part of the movie."
8. "I'd like to be related to the movie itself rather than just watching it if I use head-mounted display device."

While some comments mentioned they did not want to be distracted:

1. "If it was a movie I do not like my real body appear in the movie, it distracts me, and I cannot concentrate on the movie."
2. "Seeing my body /a virtual costume distracts me from watching the movie. I would prefer just to sit and watch the movie especially if I am not able to interact with it."
3. "Don't want to be disturbed."
4. "I thought with the costume was so cool, but would rather not see myself than see myself with no costume."

All participants were also questioned about the overall experience during the experiment, below are some of the comments from them.

1. "I like the virtual world it's fun, and I will be great to happen in video games."
2. "360-degree experience is very good."
3. "I like all of it, it was enjoyable."
4. "I think the equipment is not comfortable, maybe, I felt more uncomfortable because of my glasses."
5. "It's a fun experience to have your body with the costume in the movie."
6. "Like the real body in the movie- feel I am in the movie."
7. "I didn't like that I can see myself inside. It distracts me from watching the movie."

8. "I like now if felt I was inside the scene since I can look around."
9. "I like the VR headset."
10. "Seeing myself in the stormtrooper costume."
11. "knowing that maybe in the future the technology will make it more real."
12. "It was a funny experience."

All participants were also encouraged to identify an issue to improve the whole experience and give any suggestions about the experiment, below are some of the responses from them.

Some participants expected more interactions with the movie.

1. "More interaction."
2. "Better graphics for the real body or a more interactive interface that responds the correct way."

Some participants expected software or hardware improvements.

1. "Current technology can not make the seamless connection between the video and the added overlay."
2. "In my option, equipment needs to be lighter and comfortable."
3. "I recommend that virtual costumes could become more awesome, making people more fun.."
4. "Great idea, improvement hardware will improve the performance and experience."
5. "The movie's quality, the virtual costume stability."

A couple of participants mentioned about the wider movie options.

1. "It may give people more sense about the experiment if the movie has two different situations rather than only under the blue sky."
2. "Maybe change the movie after every condition."
3. "The movie's quality, the virtual costume stability."

4. "It has potential the video source could be some video feel from the real world, e.g. video from drone front camera, which can provide a semi-real experience of drive a plane (can see hands) while operating drones."

There were some general feedback from the participants as well.

1. "It was really cool have seen something like this before ."
2. "Only show the virtual costume without the body may feel better."
3. "Current technology can not make the seamless connection between the video and the added overlay."

According to the qualitative data analysis, it was found that most of the participants preferred watching a movie with the rendered context-aware virtual object augmented on the real body. Their comments above indicate that the virtual costume matching the movie content makes them feel more immersed and realistic in the movie.

However, there are different opinions for some participants. They would like to watch the movie without any interaction because they feel more comfortable to concentrate on the movie itself.

6.3.2 Qualitative Measures Summary

Based on the comments and feedback presented in previous section 6.3.1, the researcher categorized those comments with similar ideas.

The positive feedback for the enhanced visualization were focused on the new and fun experience that this experiment brought to them. They said the enhanced visualization made them feel connected to the movie, and these virtual costume made them feel realistic. The negative comments were mainly about the enhanced visualization distracted them to enjoy the movie itself, and they had high standard requirements expectations for the quality of virtual costumes, real body, movie content and tracking accuracy.

Most of the users had clearly understood and agreed on the good future of enhanced visualization, and they did have an open mind in the rapid development of AV and MR technology.

6.4 Summary

This chapter mainly described the experiment data analysis results collected by user experiment. The outputs of statistical analysis were reported, including the quantitative measure and qualitative measure. Next chapter will discuss the results found and identify limitations of this study.

Chapter 7

Discussion and Limitation

This chapter discusses the results found in Chapter 6 for the user experiment. It discusses the possible explanation of the different outcomes in the results and identifies limitations of the experiment.

7.1 Discussion

The results of the user experiment showed that there was a significant difference between the three conditions when the user watched 360° movies using a HMD. The user had an improved sense of presence partially by visualizing his/her own body with or without the context-aware virtual objects. There was a significant difference in user preference ranking between the three conditions, these came from the significant higher ranking between visualizing user's own body with or without the context-aware virtual objects and visualizing a movie only. This result agreed with the ranking results observed across conditions A (visualizing a movie only), B (visualizing user's real body in the movie), and C (visualizing user's real body with context-aware virtual objects in the movie). Users preferred an enhanced visualization by rendering their own body with context-aware virtual objects together, as 15 (62.5%) participants ranked this condition as the best one.

7.1.1 Sense of Presence

One of the main questions that this thesis sought to answer was whether the visual perception of the real body with virtual objects enhanced user's sense of presence. This subsection discusses the relevant analysis results and qualitative observation of interaction difference for three conditions.

The results did not support the hypothesis H_1 of this study. Overall, users who watched a movie with an enhanced visualization had a higher mean value, as well as a higher median value, although the Friedman test determined that the overall sense of presence not differ statistically significantly between three conditions. Condition C had the highest mean and median value, while condition A had the lowest mean and median value in general.

The overall IPQ consisted of four components which had been described in Chapter 5. There were significant effects on the general presence and spatial presence with the Friedman test between three conditions.

In general presence, there was a significant difference between condition A and condition B and between condition A and condition C. This shows that condition A, which visualized the movie only, had the less general user presence to users. There was no significant difference between condition B and condition C. However, the mean value for condition C was slightly higher than condition B. The answer was from "not at all" to "very much". Compared with condition A, users could see their own body in both condition B and condition C, which might cause a significant level difference in general presence. While the enhanced visualization with virtual objects would not affect the sense of "being there".

In spatial presence, there was a significant difference between three conditions shown. Again, a post hoc Wilcoxon signed-rank tests with Bonferroni correction was conducted to further analysis the pair-comparison. Therefore, only a significant level between condition C and condition A. Enhanced visualization showed an outstanding improvement for user's spatial presence. It had been supported by mean value as well. Although, there were no significant differences between condition A and condition B or condition B and C shown by Friedman test. The mean and median value could still indicate that the condition C had the highest mean and median value, while condition A had the lowest mean and median value. These still matched the experiment expectation, but not as much difference as predicted.

In involvement, there was no significant difference between three conditions. Different from General Presence and Spatial Presence, condition B had highest mean and median value in involvement, while condition C had a higher mean value, but the same median value. As this component contained the questions regarding the real world variables, like sound, the lagging issue caused by low FPS might affect the user's experience. The low FPS might be because of the 32-bit DepthSense SDK had been used in the project, due to the compatibility issue.

In experienced realism, there was no significant difference between three conditions as well. Condition A had the lowest median and mean value, while condition C had the highest the median value, but the same mean value with condition B. There was a slight improvement in experienced realism by enhanced visualization, but not as much as expected. This could happen because of the quality of real body and virtual armor and weapon. It seemed that the low level of FPS, which less than 30 frames caused the lagging issue in the experiment. Therefore, the lagging issue could affect the experienced realism experience as well.

7.1.2 User's preference

Another main question was user's preference for three conditions. The rankings for three conditions were collected in the post-experiment questionnaire.

Based on the Friedman test, a significant difference had been found between three conditions. Condition A had the worst mean rank and median rank value, while condition C had the best mean rank and median rank value. The post hoc Wilcoxon signed-rank tests with Bonferroni correction was shown that there was a significant difference between watching a movie only and an enhanced visualization by rendering their own body with context-aware virtual objects together. Users significantly preferred seeing their real body and virtual objects in the movie over only watching the movie. Hence, the hypothesis H_2 is confirmed.

However, this post hoc Wilcoxon signed-rank tests with Bonferroni correction did not show any significant difference between other conditions. Therefore, the hypothesis H_4 is not supported. These might be caused by the large number of second place the condition B had. Among the three conditions for ranking, condition C had the most number of best ranking, which was 15 out of 24, while condition A had the last place for 13 out of 24. It meant users had a strong preference in the first place for condition C and and last place for condition A. Another reason was a group of participants had explained they did not want to be disturbed by any enhanced visualization contained themselves, as it distracted them to engage in the movie itself.

A further study could be conducted regarding the enhanced visualization level to balance the engagement and immersion.

7.1.3 Overall Discussion

With the experiment results, the research questions defined in Chapter 1 can be answered.

Does the system improve user's sense of presence by enhanced visualization (visualizing the real body with or without real-time augmented context-aware virtual objects) in comparison to visualizing a movie only in a cinematic scene?

The results showed that the there was no significant improvement regarding the user's sense of presence, although, the general presence and spatial presence have been improved remarkably. The answer to this research question was: it was not wholly enhanced the user's sense of presence, but partially improved in the general presence and spatial presence. Regarding the involvement and experienced realism part, it required further investigation in the future study.

Does the user have a preference in visualizing real-time blending (nothing, only physical object or real-time rendering with context-aware virtual objects) in a cinematic scene?

The results showed that there was an outstanding difference in the preferred condition choice, users particularly preferred enhanced real-time rendering with context-aware virtual objects augmented on physical objects over watching the movie only in a cinematic scene. However, it do need further research to identify the improvements in the level of real-time blending.

Users significantly preferred seeing their real body and virtual objects in the movie over only watching the movie. However, no particular preference between seeing their body with or without virtual objects.

7.2 Limitations

The data gathered from this study was taken in an experimental setting. Despite efforts to reduce the confounding variables like the sound of surroundings, the temperature of the room, the shake of the building, there were still some impacts of different external variables on the experiment.

Based on the pre-experiment questionnaire, there were a large number of participants who had never watched 360° movies nor used HMD, that meant this would be their first VR experience which may have affected their choices in the experiment. It is possible that results could have been affected by factors that were difficult to control. To reduce noise from external factors, future studies will need to have a larger data sets.

Some technical limitations in the prototype system arose from the study, which should be noticed by other researchers in the future.

- The 3D models used in the project were very fundamental. It had limited the user experience because less sophisticated and realistic 3D models are specifically designed for the project.

- There was a technical limitation in preforming the frame rate running in Unity 3D when the real body and virtual objects are rendered together. This low frame rate issue might be limited by a the existing algorithm, configuration and hardwares.

- The accuracy of skeletal tracking in Microsoft Kinect 1 had a limited quality of a resolution, a narrower field of view and a less detailed full tracked skeleton compared with latest Kinect version.

- The virtual costumes' alignment with user's real body in Microsoft Kinect 1 had been limited by a current algorithm and the available hardware used in the project.

- When rendering the real body with SoftKinetic depth camera, there was a lot of noise in the real data which affected the experience of seamless connection and the realism.

- As described in Chapter 4, the "easy movie texture" asset had been chosen to import the 360° movie into Unity 3D. It was free and without watermark, but the quality of the 360° movie was decreased by this asset.

Besides the limitations identified above, the participants also proposed some improvements which could benefit the study in the future. Some of them expected more interactions with the movie, not only the real body and virtual objects but also real hand interaction, as in the research Khan [38] had done regarding hand gesture-based interaction.

This study had also made use of subjective interpretations of qualitative data. The researcher had to make an effort to reduce the bias of presentation. However, this might still be a bias in the study.

7.3 Summary

This chapter discussed the results obtained from the data analysis in Chapter 6. It showed the significant difference between three conditions in the user preference ranking, the general presence and spatial presence components in the sense of presence. Users had a improved general presence and spatial presence by enhanced visualization, and they are strongly preferred to visualizing enhanced visualization in the cinematic scene. Limitations of the study had also been identified. The following chapter will outline the conclusion of the study and possible future work to improve upon the enhanced immersive interaction.

Chapter 8

Conclusion and Future Work

In this chapter, a summary of the thesis project will be outlined in the first section, and the second section will briefly describe the future work related to the project.

8.1 Conclusion

In this thesis, the researcher studied the effects of enhanced visualization on the sense of presence, as well as the user's preference on visualization levels.

For the enhanced visualization, movie context-aware virtual objects had been used to augment the user's real body. This augmentation offers a new connection between movie and user. It had an advanced level on real body blended was generated to provide a realistic scene as these virtual objects matched the movie contents. It enhanced user's body visualization by adding relevant theme armor on top of the real body, which made AR more integrated with AV. The prototype developed in the project combined the real body of the user and the movie theme related 3D virtual models. To capture the real body, a depth sense camera SoftKinetic had been used. While another depth camera Microsoft Kinect had been used to track user's body movement via skeletal tracking.

The prototype was used for a user experiment in which the researcher studied the enhanced visualization of a 360° movie. The effect of enhanced visualization had been investigated on the sense of presence and the levels of preference. The visualization of 360° movies had three levels: visualize a movie only, visualize user's own body, and visualize user's own body with the context-aware virtual objects. There were 24 participants in the user experiment. The results showed that users had a preference in visualizing their bodies with context-aware virtual objects while watching the 360° video. The results had also indicated that the enhanced visualization of user's body with or without context-aware virtual objects had no significant effect on the sense of presence. Although, there was a significant improvement in some components of IPQ, which were General Presence and Spatial Presence, by visualizing user's body with the context-aware virtual objects.

Overall, evidences had been presented that enhanced visualization which blended user's body with context-aware virtual objects had a strong preference from the users. Despite the limitation of an enhanced visualization, this user study had presented a practical way of enhancing user's experience in an immersive cinematic environment.

8.2 Future Work

Although the results of the evaluation demonstrates the users' preference for enhanced visualization of user body with the context-aware virtual objects from users, as well as the improvement in General Presence and Spatial Presence, there is need for future work to improve upon the prototype presented in the thesis. Future direction will be to improve the prototype based on the limitations described in Chapter 7.

Some of the possibilities are as followed.

- As described in the Chapter 7 limitation section, there are some participants who have never used HMD and watched 360° movies. In the future study, one more independent variable can be added regarding user's skills level of VR, AR, and MR, like "beginner" or "expert". Another user study can be conducted based on two independent variables.

- The user study in this thesis has limited virtual costumes to the real body. Other relevant 3D-model costumes can be added, and a user experiment can be carried out to explore the effect of changing costumes on the sense of presence.

- A between-subject user study can be conducted if multiple 360° movies and virtual costumes are added to the project. Researchers can prepare more than one set of 360° movies, contains different topics, like video from a drone front camera, simulate the driving experience for the user, to explore whether these improvements can affect user's sense of presence.

- Based on some of the user's feedback after the user experiment, more interaction with virtual weapon or armor can be integrated, like hand gesture-based interaction with a virtual weapon. A user study can be conducted with these interactions integrated into the current prototype, to investigate their effect on the sense of presence.

- On the technical compatibility aspect, it will be better to solve the 32-bit and 64-bit compatibility issue between Softkinetic camera, Kinect, and Unity 3D. This improvement may increase the FPS in the user experiment, which can affect user's experience.

- Another technical aspect that can be improved in the prototype is the optimization to reduce the crashes caused by memory leak. As the SoftKinetic Depthsense SDK only supports Windows 7, a hardware replacement which supports Windows 10 will be better to use in the future.

This thesis is just an initial exploration in the augmented virtuality aspect of the immersive cinematic environment. The project has considerable possibilities to improve and optimize in both hardware and software aspects. There is a great potential in the AV enhanced visualization in an cinematic environment by integrating more interactions and features in the future.

Bibliography

[1] Eunji Oh, Minkyoung Lee, and Sujin Lee. How 4d effects cause different types of presence experience? In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, pages 375–378. ACM, 2011.

[2] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for industrial applications*, pages 282–292. International Society for Optics and Photonics, 1995.

[3] Paul Milgram and Fumio Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[4] Christopher Stapleton, Charles Hughes, Michael Moshell, Paulius Micikevicius, and Marty Altman. Applying mixed reality to entertainment. *Computer*, 35(12):122–124, 2002.

[5] Joshua Yang Chen. Augmenting immersive cinematic experience/scene with user body visualization. 2016.

[6] Sony. *Depth-Market-HMD*, 2018 (accessed Feb 22, 2018). URL `https://www.sony-depthsensing.com/Depthsense/Markets/HMD`.

[7] Sang Chul Ahn, Ig-Jae Kim, Hyoung-Gon Kim, Yong-Moo Kwon, and Heedong Ko. Audience interaction for virtual reality theater and its implementation. *ACM International Symposium on Virtual Reality Software and Technology (VRST 2001)*, pages 41–45, 2001.

[8] ChangHoon Park, Heedong Ko, and Taiyun Kim. Naver: Networked and augmented virtual environment architecture; design and implementation of vr framework for gyeongju vr theater. *Computers Graphics*, 27(2):223–230, 2003.

[9] Adrian David Cheok, Wang Weihua, Xubo Yang, Simon Prince, Fong Siew Wan, Mark Billinghurst, and Hirokazu Kato. Interactive theatre experience in embodied+

wearable mixed reality space. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality*, page 59. IEEE Computer Society, 2002.

[10] Framestore Studio. *Interstellar Virtual Reality Experience*, 2016 (accessed June 30, 2016). URL `http://framestorevr.com/interstellar-experience/`.

[11] Hideyuki Tamura, Hiroyuki Yamamoto, and Akihiro Katayama. Mixed reality: Future dreams seen at the border between real and virtual worlds. *Computer Graphics and Applications, IEEE*, 21(6):64–70, 2001.

[12] Joshua Chen, Gun A. Lee, Mark Billinghurst, Robert W. Lindeman, and Christoph Bartneck. The Effect of User Embodiment in AV Cinematic Experience. In Robert W. Lindeman, Gerd Bruder, and Daisuke Iwai, editors, *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. The Eurographics Association, 2017. ISBN 978-3-03868-038-3. doi: 10.2312/egve.20171339.

[13] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. The magicbook: a transitional ar interface. *Computers Graphics*, 25(5):745–753, 2001.

[14] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. The magicbook-moving seamlessly between reality and virtuality. *IEEE Computer Graphics and applications*, 21(3):6–8, 2001.

[15] A Duenser, M Billinghurst, and R Grasset. Moving between contexts-a user evaluation of a transitional interface. *ICAT,*, 2008.

[16] Humayun Khan, Gun A. Lee, Simon Hoermann, Rory M. S. Clifford, Mark Billinghurst, and Robert W. Lindeman. Evaluating the Effects of Hand-gesture-based Interaction with Virtual Content in a 360 Movie. In Robert W. Lindeman, Gerd Bruder, and Daisuke Iwai, editors, *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments*. The Eurographics Association, 2017. ISBN 978-3-03868-038-3. doi: 10.2312/egve.20171340.

[17] Zhengyou Zhang. Microsoft kinect sensor and its effect. *IEEE multimedia*, 19(2): 4–10, 2012.

[18] Jan Smisek, Michal Jancosek, and Tomas Pajdla. 3d with kinect. *Proc.IEEE ICCV Workshops*, pages 1154–1160, 2011.

[19] Microsoft. *Skeletal Tracking*, 2017 (accessed Jan 6, 2017). URL `https://msdn.microsoft.com/en-us/library/hh973074.aspx`.

[20] Microsoft. *Joint Orientation*, 2017 (accessed Jan 6, 2017). URL `https://msdn.microsoft.com/en-us/library/hh973073.aspx`.

[21] Lucía Vera, Jesús Gimeno, Inmaculada Coma, and Marcos Fernández. Augmented mirror: interactive augmented reality system based on kinect. In *IFIP Conference on Human-Computer Interaction*, pages 483–486. Springer, 2011.

[22] Dimitrios S Alexiadis, Philip Kelly, Petros Daras, Noel E O'Connor, Tamy Boubekeur, and Maher Ben Moussa. Evaluating a dancer's performance using kinect-based skeleton tracking. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 659–662. ACM, 2011.

[23] Valentino Frati and Domenico Prattichizzo. Using kinect for hand tracking and rendering in wearable haptics. In *World Haptics Conference (WHC), 2011 IEEE*, pages 317–321. IEEE, 2011.

[24] Adrian Clark and Thammathip Piumsomboon. A realistic augmented reality racing game using a depth-sensing camera. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, pages 499–502. ACM, 2011.

[25] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohi, Jamie Shotton, Steve Hodges, and Andrew Fitzgibbon. Kinectfusion: Real-time dense surface mapping and tracking. In *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*, pages 127–136. IEEE, 2011.

[26] Shahram Izadi, David Kim, Otmar Hilliges, David Molyneaux, Richard Newcombe, Pushmeet Kohli, Jamie Shotton, Steve Hodges, Dustin Freeman, Andrew Davison, et al. Kinectfusion: real-time 3d reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 559–568. ACM, 2011.

[27] Leandro Cruz, Djalma Lucio, and Luiz Velho. Kinect and rgbd images: Challenges and applications. In *Graphics, Patterns and Images Tutorials (SIBGRAPI-T), 2012 25th SIBGRAPI Conference on*, pages 36–49. IEEE, 2012.

[28] Jim Blascovich. Social influence within immersive virtual environments. In *The social life of avatars*, pages 127–145. Springer, 2002.

[29] Kairat Aitpayev and Jaafar Gaber. Creation of 3d human avatar using kinect. *Asian Transactions on Fundamentals of Electronics, Communication & Multimedia*, 1(5): 12–24, 2012.

[30] Thibaut Weise, Sofien Bouaziz, Hao Li, and Mark Pauly. Realtime performance-based facial animation. In *ACM transactions on graphics (TOG)*, volume 30, page 77. ACM, 2011.

[31] Omprakash Rudhru. Visual affordances in natural user interfaces. 2016.

[32] Björn Nilson and Martin Söderberg. Game engine architecture. *Mälardalen University*, 2007.

[33] Mohd Fairuz Shiratuddin and Walid Thabet. Virtual office walkthrough using a 3d game engine. *International Journal of Design Computing*, 4, 2002.

[34] PD Bourke. idome: Immersive gaming with the unity game engine. *Proceedings of the Computer Games & Allied Technology*, 9:265–272, 2009.

[35] Antônio Carlos A. Mól, Carlos Alexandre F. Jorge, and Pedro M. Couto. Using a game engine for vr simulations in evacuation planning. *Computer Graphics and Applications, IEEE*, 28(3):6–12, 2008.

[36] Hrvoje Benko, Ricardo Jota, and Andrew Wilson. Miragetable: freehand interaction on a projected augmented reality tabletop. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 199–208. ACM, 2012.

[37] Unity Technologies. *Unity - Game Engine*, 2016 (accessed June 29, 2016). URL `https://unity3d.com/`.

[38] Humayun Khan. Hand gesture-based interaction in an immersive cinematic environment. 2017.

[39] Microsoft. *Kinect Fusion*, 2017 (accessed Jan 11, 2017). URL `https://msdn.microsoft.com/en-us/library/dn188670.aspx`.

[40] Universtiy of Canterbury Human Interface Technology Laboratory New Zealand. Project midterm progress report: February june 2016 human-digital content interaction for immersive 4d home entertainment. Report, University of Canterbury, June 2016.

[41] Tobias Günther, Ingmar S Franke, and Rainer Groh. Aughanded virtuality-the hands in the virtual environment. In *3D User Interfaces (3DUI), 2015 IEEE Symposium on*, pages 157–158. IEEE, 2015.

[42] Zugara Inc. *How Does The Kinect 2 Compare To The Kinect 1?*, 2017 (accessed Jan 9, 2017). URL `http://zugara.com/how-does-the-kinect-2-compare-to-the-kinect-1`.

[43] SoftKinetic. *DS325 Datasheet*, (accessed: 10.01.2017). URL http://www.softkinetic.com/Portals/0/Download/WEB_20120907_SK_DS325_Datasheet_V2.1.pdf.

[44] LLC Oculus VR. *DK2,Development Kit 2*, 2017 (accessed Jan 10, 2017). URL https://www3.oculus.com/en-us/dk2/.

[45] LLC Oculus VR. *Overview*, 2017 (accessed Jan 10, 2017). URL https://www3.oculus.com/en-us/rift/.

[46] Unity Technologies. *About us*, 2017 (accessed Jan 11, 2017). URL https://unity3d.com/public-relations.

[47] Unity Technologies. *VR overview*, 2017 (accessed Jan 11, 2017). URL https://docs.unity3d.com/Manual/VROverview.html.

[48] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research Methods in Human-Computer Interaction*. John Wiley & Sons, 2010.

[49] Thomas Schubert, Frank Friedmann, and Holger Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001.

Appendix A

Information Sheet and Consent Form

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

HUMAN ETHICS COMMITTEE

Secretary, Rebecca Robinson
Telephone: +64 03 369 4588, Extn 94588
Email: human-ethics@canterbury.ac.nz

Ref: HEC 2017/04/LR

7 February 2017

Wenjing Tang
HITLab NZ
UNIVERSITY OF CANTERBURY

Dear Wenjing

Thank you for submitting your low risk application to the Human Ethics Committee for the research proposal titled "Augmented Virtuality in an Immersive Cinematic".

I am pleased to advise that the application has been reviewed and approved.

Please note that this approval is subject to the incorporation of the amendments you have provided in your emails of 25th and 31st January 2017.

With best wishes for your project.

Yours sincerely

pp. R. Robinson

Associate Professor Jane Maidment
Chair, Human Ethics Committee

UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Department: HIT Lab NZ
Telephone: +64 3 364 2349
Email: jady.tang@pg.canterbury.ac.nz
31.12.2016

Augmented Virtuality in an Immersive Cinematic
Information Sheet for Experiment Participants

RESEARCHER: Wenjing Tang (jady.tang@pg.canterbury.ac.nz)

SUPERVISORS: Prof. Rob Lindeman (gogo@hitlabnz.org), Dr. Gun Lee (gun.lee@canterbury.ac.nz),

Prof. Mark Billinghurst (mark.billinghurst@canterbury.ac.nz)

INTRODUCTION:

You are invited to take part in a cinematic experience research study. Before you decide to be part of this study, you need to understand the risks and benefits. This information sheet provides information about the research study. The researcher will be available to answer your questions and provide further explanations. If you agree to take part in the research study, you will be asked to sign the consent form.

PURPOSE:

The purpose of this study to identify the effect of Augmented Reality and Augmented Virtuality on user experience while playing cinematic content through a head-mounted display (HMD) such as the Oculus Rift.

PROCEDURES:

If you choose to take part in this study, your involvement in this project, the experimental procedure is outlined as below:

a. The participant reads the information sheet and signs the consent form.

b. The participant answers to a pre-experiment questionnaire individually about demographic information and his/her previous experience with using HMD.

c. The researcher explains the study setup and experimental tasks for the participant to perform during the study.

d. The participant performs the experimental tasks including:
- Wearing the Oculus Rift HMD and pilot testing with mounted camera
- Perform the given task, which may include watching a short 360 Panorama movie, either with or without one's own body blended into the movie or augmenting with the virtual objects
- Rate personal thoughts and feelings based on each task by answering a questionnaire.

e. The participant answers a post-experiment questionnaire asking for feedback on the overall study.

The whole procedure will take approximately 35-40 minutes.

Weniing Tang

RISKS/DISCOMFORTS:

Risks are dependent on individuals in this study. As you will be wearing a head-mounted display (HMD) a number of times, you might feel some form of nausea or giddiness. Please tell the researcher when you feel uncomfortable. Being still with eyes closed could help further prevent nausea coming in rapidly.

Participation is voluntary and you have the right to withdraw at any stage without penalty

CONFIDENTIALITY:

All data obtained from participants will be kept confidential. In publications (e.g.
Thesis, a public document which will be available through the UC Library), we will mainly report the results in an aggregate format: reporting only combined results and never reporting individual ones. In case of reporting quotes of the participants from the questionnaires, we will keep the source anonymous. Video of the experiment will be recorded for analysis purposes. The recorded video will be only of what participants can see through the Oculus Rift head mounted display. There will be no recordings of the participant's face, maintaining the anonymity of the participants. All recordings will be concealed, and none other than the researchers will have access to them. The data will be kept securely for a minimum period of 5 years and will be destroyed after completion of the research project.

Please indicate to the researcher on the consent form if you would like to receive a copy of the summaryof results of the project.

The project is being carried out *Master of Human Interface Technology* by Wenjing Tang under the supervision of supervisors list above, who can be contacted at the emails. They will be pleased to discuss any concerns you may have about participation in the project.

This project has been reviewed and approved by the University of Canterbury Human Ethics Committee, and participants should address any complaints to The Chair, Human Ethics Committee, University of Canterbury, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz).

If you agree to participate in the study, you are asked to complete the consent form.

UC
UNIVERSITY OF CANTERBURY
Te Whare Wānanga o Waitaha
CHRISTCHURCH NEW ZEALAND

Weniing Tang

Department: HIT Lab NZ
Telephone: +64 3 364 2349
Email: jady.tang@pg.canterbury.ac.nz

Augmented Virtuality in an Immersive Cinematic
Consent Form for Experiment Participants

RESEARCHER: Wenjing Tang (jady.tang@pg.canterbury.ac.nz)

SUPERVISORS: Prof. Rob Lindeman (gogo@hitlabnz.org), Dr. Gun Lee (gun.lee@canterbury.ac.nz),

Prof. Mark Billinghurst (mark.billinghurst@canterbury.ac.nz)

Include a statement regarding each of the following:

- ☐ I have been given a full explanation of this project and have had the opportunity to ask questions.
- ☐ I understand what is required of me if I agree to take part in the research.
- ☐ I understand that participation is voluntary and I may withdraw at any time without penalty. Withdrawal of participation will also include the withdrawal of any information I have provided should this remain practically achievable.
- ☐ I understand that any information or opinions I provide will be kept confidential to the researcher administrators of the research project and that any published or reported results will not identify the participants. I understand that a thesis is a public document and will be available through the UC Library.
- ☐ I understand that all data collected for the study will be kept in locked and secure facilities and/or in password protected electronic form and will be destroyed after *five* years.
- ☐ I understand the risks associated with taking part and how they will be managed.
- ☐ I understand that I can contact the researcher or supervisors list above for further information. If I have any complaints, I can contact the Chair of the University of Canterbury Human Ethics Committee, Private Bag 4800, Christchurch (human-ethics@canterbury.ac.nz)
- ☐ I would like a summary of the results of the project.
- ☐ By signing below, I agree to participate in this research project, and I authorize recordings or other materials taken from this study used for scientific purpose, and I consent to publication of the results of the study.

______________________________	______________________________	______________
Participant (Print name)	Signature	Date

Email address *(for report of findings, if applicable)*:____

Wenjing Tang

Appendix B

Questionnaires

Pre-experiment Questionnaire

To be filled by researcher:

Participant No.

1. What is your age?

   ______ years old

2. What is your gender?

   ❍ Female ❍ Male ❍ Other

3. Have you used Head-Mounted Displays before?
   (For example, Oculus Rift, Samsung Gear VR, or HTC Vive)

   ❍ Never
   ❍ Few times a year
   ❍ Few times a month
   ❍ Few times a week
   ❍ Everyday

4. Have you watched a 360 movie on Head-Mounted Displays before?
   (For example, Oculus Rift, Samsung Gear VR, or HTC Vive)

   ❍ Never
   ❍ Few times a year
   ❍ Few times a month
   ❍ Few times a week
   ❍ Everyday

5. Would like to see my real body appearing in the movie

Fully disagree			Neutral			Fully agree
○	○	○	○	○	○	○

6. Please explain why you rated as above:

7. I would like to see my real body with virtual costume on matching the movie theme.

Fully disagree			Neutral			Fully agree
○	○	○	○	○	○	○

8. Please explain why you rated as above:

Thank you !

Per-condition Questionnaire

To be filled by researcher

Participant No.	Condition No.

INFO: For questions below please choose how much you agree to each statement below

Q1. How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?

Extremely aware — Moderately aware — Not aware at all

○ ○ ○ ○ ○ ○ ○

Q2. How real did the virtual world seem to you?

Completely real — Not real at all

○ ○ ○ ○ ○ ○ ○

Q3. I had a sense of acting in the virtual space, rather than operating something from outside.

Fully disagree — Fully agree

○ ○ ○ ○ ○ ○ ○

Q4. How much did your experience in the virtual environment seem consistent with your real world experience?

Not consistent — Moderately consistent — Very consistent

○ ○ ○ ○ ○ ○ ○

Q5. How real did the virtual world seem to you?

About as real as an imagined world — Indistinguishable from the real world

○ ○ ○ ○ ○ ○ ○

Q6. I did not feel present in the virtual space.

Did not feel — Felt present

○ ○ ○ ○ ○ ○ ○

Q7. I was not aware of my real environment.

Fully disagree — Fullyagree

○ ○ ○ ○ ○ ○ ○

Q8. In the computer generated world I had a sense of “being there”

Not at all very much

○ ○ ○ ○ ○ ○ ○

Q9. Somehow I felt that the virtual world surrounded me.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q10. I felt present in the virtual space.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q11. I still paid attention to the real environment.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q12. The virtual world seemed more realistic than the real world.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q13. I felt like I just perceiving pictures.

Fully disagree fully agree

○ ○ ○ ○ ○ ○ ○

Q14. I was completely captivated by the virtual world.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q15. I lost myself in the cinematic experience.

Fully disagree Fully agree

○ ○ ○ ○ ○ ○ ○

Q16. This cinematic experience was fun.

Fully disagree Fullyagree

Q17. Without regarding the interface used to watch, I like the content of the movie.

Fully disagree Fullyagree

Thank you!

Post-experiment Questionnaire

To be filled by researcher:

Participant No.

Q1. I like having my real body appear in the movie.

Fully disagree ○ ○ ○ ○ ○ ○ ○ Fully agree

Q2. I like having my real body with virtual costume on matching the movie theme.

Fully disagree ○ ○ ○ ○ ○ ○ ○ Fully agree

Q3. Please rank the three conditions based on your preference by numbering (1: best ~ 3: worst) them.

___ Not seeing the body

___ Seeing my real body without virtual costume

___ Seeing my real body with virtual costume

Q4. Please explain why you choose as above in the previous two questions.

Q5. What did you like overall?

Q6. What did you dislike overall?

Q7. What could be improved?

Q8. Any other comments or suggestions on the experiment?

Thank you!